



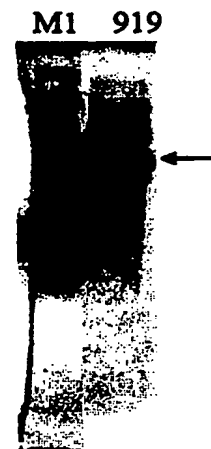
## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/31, C07K 14/22, 16/12, C12Q 1/68, A61K 39/095, G01N 33/50</b>		<b>A2</b>	(11) International Publication Number: <b>WO 99/57280</b>	
			(43) International Publication Date: 11 November 1999 (11.11.99)	
(21) International Application Number: <b>PCT/US99/09346</b>		(US). MASIGNANI, Vega [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). MORA, Marirosa [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). PETERSEN, Jeremy [US/US]; Arlington, VA (US). PIZZA, Mariagratia [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RAPPUOLI, Rino [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RATTI, Giulio [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCALATO, Enzo [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCARSELLI, Maria [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). TETTELIN, Herve [US/US]; Gaithersburg, MD (US). VENTER, J., Craig [US/US]; Rockville, MD (US).		
(22) International Filing Date: 30 April 1999 (30.04.99)				
(30) Priority Data:				
60/083,758	1 May 1998 (01.05.98)			US
60/094,869	31 July 1998 (31.07.98)			US
60/098,994	2 September 1998 (02.09.98)	US		
60/099,062	2 September 1998 (02.09.98)	US		
60/103,749	9 October 1998 (09.10.98)	US		
60/103,794	9 October 1998 (09.10.98)	US		
60/103,796	9 October 1998 (09.10.98)	US		
60/121,528	25 February 1999 (25.02.99)	US		
(71) Applicants (for all designated States except US): CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). THE INSTITUTE FOR GENOMIC RESEARCH [US/US]; 9212 Medical Center Drive, Rockville, MD 20850 (US).		(74) Agent: HARBIN, Alisa, A.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).		
(72) Inventors; and				
(75) Inventors/Applicants (for US only): FRASER, Claire [US/US]; Rockville, MD (US). GALEOTTI, Cesira [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). GRANDI, Guido [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). HICKEY, Erin [US/US]; Gaithersburg, MD				
		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).		
		<b>Published</b> Without international search report and to be republished upon receipt of that report.		

(54) Title: NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

## (57) Abstract

The invention provides proteins from Neisseria meningitidis, including the amino acid sequences and the corresponding nucleotide sequences. The proteins are predicted to be useful antigens for vaccines and/or diagnostics.

**919 (46 kDa)****A) PURIFICATION**

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

This application is a continuation-in-part of the following U.S. Provisional Patent applications, from each of which priority is claimed, and each of which is incorporated by reference in its entirety: 60/083,758 (filed May 1, 1998); 60/094,869 (filed July 31, 1998); 60/098,994 (filed September 2, 1998); 60/099,062 (filed September 2, 1998); 60/103,749 (filed October 9, 1998); 60/103,794 (filed October 9, 1998); 60/103,796 (filed October 9, 1998); and 60/121,528 (filed February 25, 1999).

This invention relates to antigens from the bacterial species: *Neisseria meningitidis* and *Neisseria gonorrhoeae*.

### BACKGROUND

*Neisseria meningitidis* is a non-motile, gram negative diplococcus human pathogen. It colonizes the pharynx, causing meningitis and, occasionally, septicaemia in the absence of meningitis. It is closely related to *N. gonorrhoea*, although one feature that clearly differentiates meningococcus from gonococcus is the presence of a polysaccharide capsule that is present in all pathogenic meningococci.

*N. meningitidis* causes both endemic and epidemic disease. In the United States the attack rate is 0.6-1 per 100,000 persons per year, and it can be much greater during outbreaks. (see Lieberman *et al.* (1996) Safety and Immunogenicity of a Serogroups A/C *Neisseria meningitidis* Oligosaccharide-Protein Conjugate Vaccine in Young Children. *JAMA* 275(19):1499-1503; Schuchat *et al* (1997) Bacterial Meningitis in the United States in 1995. *N Engl J Med* 337(14):970-976). In developing countries, endemic disease rates are much higher and during epidemics incidence rates can reach 500 cases per 100,000 persons per year. Mortality is extremely high, at 10-20% in the United States, and much higher in developing countries. Following the introduction of the conjugate vaccine against *Haemophilus influenzae*, *N. meningitidis* is the major cause of bacterial meningitis at all ages in the United States (Schuchat *et al* (1997) *supra*).

Based on the organism's capsular polysaccharide, 12 serogroups of *N. meningitidis* have been identified. Group A is the pathogen most often implicated in epidemic disease in sub-Saharan Africa. Serogroups B and C are responsible for the vast majority of cases in the

United States and in most developed countries. Serogroups W135 and Y are responsible for the rest of the cases in the United States and developed countries. The meningococcal vaccine currently in use is a tetravalent polysaccharide vaccine composed of serogroups A, C, Y and W135. Although efficacious in adolescents and adults, it induces a poor immune response and short duration of protection, and cannot be used in infants [eg. Morbidity and Mortality weekly report, Vol.46, No. RR-5 (1997)]. This is because polysaccharides are T-cell independent antigens that induce a weak immune response that cannot be boosted by repeated immunization. Following the success of the vaccination against *H.influenzae*, conjugate vaccines against serogroups A and C have been developed and are at the final stage of clinical testing (Zollinger WD "New and Improved Vaccines Against Meningococcal Disease". In: *New Generation Vaccines, supra*, pp. 469-488; Lieberman *et al* (1996) *supra*; Costantino *et al* (1992) Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. *Vaccine* 10:691-698).

Meningococcus B (menB) remains a problem, however. This serotype currently is responsible for approximately 50% of total meningitis in the United States, Europe, and South America. The polysaccharide approach cannot be used because the menB capsular polysaccharide is a polymer of  $\alpha(2-8)$ -linked *N*-acetyl neuraminic acid that is also present in mammalian tissue. This results in tolerance to the antigen; indeed, if an immune response were elicited, it would be anti-self, and therefore undesirable. In order to avoid induction of autoimmunity and to induce a protective immune response, the capsular polysaccharide has, for instance, been chemically modified substituting the *N*-acetyl groups with *N*-propionyl groups, leaving the specific antigenicity unaltered (Romero & Outschoorn (1994) Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? *Clin Microbiol Rev* 7(4):559-575).

Alternative approaches to menB vaccines have used complex mixtures of outer membrane proteins (OMPs), containing either the OMPs alone, or OMPs enriched in porins, or deleted of the class 4 OMPs that are believed to induce antibodies that block bactericidal activity. This approach produces vaccines that are not well characterized. They are able to protect against the homologous strain, but are not effective at large where there are many antigenic variants of the outer membrane proteins. To overcome the antigenic variability, multivalent vaccines containing up to nine different porins have been constructed (eg. Poolman JT (1992) Development of a meningococcal vaccine. *Infect. Agents Dis.* 4:13-28).



Additional proteins to be used in outer membrane vaccines have been the opa and opc proteins, but none of these approaches have been able to overcome the antigenic variability (eg. Ala'Aldeen & Borriello (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. *Vaccine* 14(1):49-53).

A certain amount of sequence data is available for meningococcal and gonococcal genes and proteins (e.g. EP-A-0467714, WO96/29412), but this is by no means complete. Other men B proteins may include those listed in WO 97/28273, WO 96/29412, WO 95/03413, US 5,439,808, and US 5,879,686.

The provision of further sequences could provide an opportunity to identify secreted or surface-exposed proteins that are presumed targets for the immune system and which are not antigenically variable. For instance, some of the identified proteins could be components of efficacious vaccines against meningococcus B, some could be components of vaccines against all meningococcal serotypes, and others could be components of vaccines against all pathogenic *Neisseriae* including *Neisseria meningitidis* or *Neisseria gonorrhoeae*. Those sequences specific to *N. meningitidis* or *N. gonorrhoeae* that are more highly conserved are further preferred sequences.

It is thus an object of the invention is to provide Neisserial DNA sequences which encode proteins that are antigenic or immunogenic.

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the products of protein expression and purification of the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 2 illustrates the products of protein expression and purification of the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 3 illustrates the products of protein expression and purification of the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 4 illustrates the products of protein expression and purification of the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 5 illustrates the products of protein expression and purification of the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 6 illustrates the products of protein expression and purification of the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 7 illustrates the products of protein expression and purification of the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 8 illustrates the products of protein expression and purification of the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 9 illustrates the products of protein expression and purification of the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 10 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 11 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 12 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 13 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 14 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 15 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 16 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 17 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 18 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 19 shows an alignment comparison of amino acid sequences for ORF 225 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 20 shows an alignment comparison of amino acid sequences for ORF 235 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 21 shows an alignment comparison of amino acid sequences for ORF 287 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 22 shows an alignment comparison of amino acid sequences for ORF 519 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 23 shows an alignment comparison of amino acid sequences for ORF 919 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

## THE INVENTION

The invention provides proteins comprising the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples.

It also provides proteins comprising sequences homologous (i.e., those having sequence identity) to the *N. meningitidis* amino acid sequences disclosed in the examples. Depending on the particular sequence, the degree of homology (sequence identity) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). These proteins include mutants and allelic variants of the sequences disclosed in the examples. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman

homology search algorithm as implemented in MPSRCH program (Oxford Molecular) using an affine gap search with parameters: gap penalty 12, gap extension penalty 1.

The invention further provides proteins comprising fragments of the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples. The fragments should comprise at least  $n$  consecutive amino acids from the sequences and, depending on the particular sequence,  $n$  is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20 or more). Preferably the fragments comprise an epitope from the sequence.

The proteins of the invention can, of course, be prepared by various means (eg. recombinant expression, purification from cell culture, chemical synthesis *etc.*) and in various forms (eg. native, fusions *etc.*). They are preferably prepared in substantially pure or isolated form (*ie.* substantially free from other *N. meningitidis* or *N. gonorrhoeae* host cell proteins).

According to a further aspect, the invention provides antibodies which bind to these proteins. These may be polyclonal or monoclonal and may be produced by any suitable means.

According to a further aspect, the invention provides nucleic acid comprising the *N. meningitidis* nucleotide sequences and *N. gonorrhoeae* nucleotide sequences disclosed in the examples.

According to a further aspect, the invention comprises nucleic acids having sequence identity of greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) to the nucleic acid sequences herein. Sequence identity is determined as above-discussed.

According to a further aspect, the invention comprises nucleic acid that hybridizes to the sequences provided herein. Conditions for hybridization are set forth herein.

Nucleic acid comprising fragments of these sequences are also provided. These should comprise at least  $n$  consecutive nucleotides from the *N. meningitidis* sequences or *N. gonorrhoeae* sequences and depending on the particular sequence,  $n$  is 10 or more (eg 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

According to a further aspect, the invention provides nucleic acid encoding the proteins and protein fragments of the invention.

It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (eg. for antisense or probing purposes).

Nucleic acid according to the invention can, of course, be prepared in many ways (eg. by chemical synthesis, in part or in whole, from genomic or cDNA libraries, from the

organism itself *etc.*) and can take various forms (*eg.* single stranded, double stranded, vectors, probes *etc.*).

In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also protein nucleic acids (PNA) *etc.*

According to a further aspect, the invention provides vectors comprising nucleotide sequences of the invention (*eg.* expression vectors) and host cells transformed with such vectors.

According to a further aspect, the invention provides compositions comprising protein, antibody, and/or nucleic acid according to the invention. These compositions may be suitable as vaccines, for instance, or as diagnostic reagents or as immunogenic compositions.

The invention also provides nucleic acid, protein, or antibody according to the invention for use as medicaments (*eg.* as vaccines) or as diagnostic reagents. It also provides the use of nucleic acid, protein, or antibody according to the invention in the manufacture of (I) a medicament for treating or preventing infection due to Neisserial bacteria (ii) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria or (iii) for raising antibodies. Said Neisserial bacteria may be any species or strain (such as *N. gonorrhoeae*) but are preferably *N. meningitidis*, especially strain B or strain C.

The invention also provides a method of treating a patient, comprising administering to the patient a therapeutically effective amount of nucleic acid, protein, and/or antibody according to the invention.

According to further aspects, the invention provides various processes.

A process for producing proteins of the invention is provided, comprising the step of culturing a host cell according to the invention under conditions which induce protein expression.

A process for detecting polynucleotides of the invention is provided, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting said duplexes.

A process for detecting proteins of the invention is provided, comprising the steps of: (a) contacting an antibody according to the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complexes; and (b) detecting said complexes.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

Methodology - Summary of standard procedures and techniques.

General

This invention provides *Neisseria meningitidis* menB nucleotide sequences, amino acid sequences encoded therein. With these disclosed sequences, nucleic acid probe assays and expression cassettes and vectors can be produced. The expression vectors can be transformed into host cells to produce proteins. The purified or isolated polypeptides (which may also be chemically synthesized) can be used to produce antibodies to detect menB proteins. Also, the host cells or extracts can be utilized for biological assays to isolate agonists or antagonists. In addition, with these sequences one can search to identify open reading frames and identify amino acid sequences. The proteins may also be used in immunogenic compositions, antigenic compositions and as vaccine components.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature e.g., Sambrook *Molecular Cloning: A Laboratory Manual, Second Edition* (1989); *DNA Cloning, Volumes I and ii* (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); *Transcription and Translation* (B.D. Hames & S.J. Higgins eds. 1984); *Animal Cell Culture* (R.I. Freshney ed. 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); the *Methods in Enzymology* series (Academic Press, Inc.), especially volumes 154 & 155; *Gene Transfer Vectors for Mammalian Cells* (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer and Walker, eds. (1987), *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); Scopes, (1987) *Protein Purification: Principles and Practice*, Second Edition (Springer-Verlag, N.Y.), and *Handbook of Experimental Immunology, Volumes I-IV* (D.M. Weir and C.C. Blackwell eds 1986).

Standard abbreviations for nucleotides and amino acids are used in this specification.

All publications, patents, and patent applications cited herein are incorporated in full by reference.

### Expression systems

The *Neisseria menB* nucleotide sequences can be expressed in a variety of different expression systems; for example those used with mammalian cells, plant cells, baculoviruses, bacteria, and yeast.

#### i. Mammalian Systems

Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25-30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation (Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian Cells." In *Molecular Cloning: A Laboratory Manual*, 2nd ed.).

Mammalian viral genes are often highly expressed and have a broad host range; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, also provide useful promoter sequences. Expression may be either constitutive or regulated (inducible). Depending on the promoter selected, many promoters may be inducible using known substrates, such as the use of the mouse mammary tumor virus (MMTV) promoter with the glucocorticoid responsive element (GRE) that is induced by glucocorticoid in hormone-responsive transformed cells (see for example, U.S. Patent 5,783,681).

The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a

regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter (Maniatis et al. (1987) *Science* 236:1237; Alberts et al. (1989) *Molecular Biology of the Cell*, 2nd ed.). Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer (Dijkema et al (1985) *EMBO J.* 4:761) and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus (Gorman et al. (1982b) *Proc. Natl. Acad. Sci.* 79:6777) and from human cytomegalovirus (Boshart et al. (1985) *Cell* 41:521). Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion (Sassone-Corsi and Borelli (1986) *Trends Genet.* 2:215; Maniatis et al. (1987) *Science* 236:1237).

A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus tripartite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation (Birnstiel et al. (1985) *Cell* 41:349; Proudfoot and Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. In *Transcription and splicing* (ed. B.D. Hames and D.M.



Glover); Proudfoot (1989) *Trends Biochem. Sci.* 14:105). These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator/polyadenylation signals include those derived from SV40 (Sambrook et al (1989) "Expression of cloned genes in cultured mammalian cells." In *Molecular Cloning: A Laboratory Manual*).

Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviruses, such as SV40 (Gluzman (1981) *Cell* 23:175) or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein-Barr virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a prokaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 (Kaufman et al. (1989) *Mol. Cell. Biol.* 9:946) and pHEBO (Shimizu et al. (1986) *Mol. Cell. Biol.* 6:1074).

The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and a number of other cell lines.

## ii. Plant Cellular Expression Systems

There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: U.S. 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, *Phytochemistry* 30:3861-3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler et al., *Plant Molecular Biology* 3:407-418 (1984); Rogers, *J. Biol. Chem.* 260:3731-3738 (1985); Rothstein et al., *Gene* 55:353-356 (1987); Whittier et al., *Nucleic Acids Research* 15:2515-2535 (1987); Wirsal et al., *Molecular Microbiology* 3:3-14 (1989); Yu et al., *Gene* 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, Gibberellins: in: *Advanced Plant Physiology*, Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, *Plant Cell*, 2:1027-1038(1990); Maas et al., *EMBO J.* 9:3447-3452 (1990); Benkel and Hickey, *Proc. Natl. Acad. Sci.* 84:1337-1339 (1987)

Typically, using techniques known in the art, a desired polynucleotide sequence is inserted into an expression cassette comprising genetic regulatory elements designed for operation in plants. The expression cassette is inserted into a desired expression vector with companion sequences upstream and downstream from the expression cassette suitable for expression in a plant host. The companion sequences will be of plasmid or viral origin and provide necessary characteristics to the vector to permit the vectors to move DNA from an original cloning host, such as bacteria, to the desired plant host. The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for *Agrobacterium* transformations, T DNA sequences for *Agrobacterium*-mediated transfer to plant chromosomes. Where the heterologous gene is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers, for example for the members of the grass family, is found in Wilmink and Dons, 1993, *Plant Mol. Biol. Repr.*, 11(2):165-185.

Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome. Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The nucleic acid molecules of the subject invention may be included into an expression cassette for expression of the protein(s) of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous protein encoding sequence the following elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

A heterologous coding sequence may be for any protein relating to the present invention. The sequence encoding the protein of interest will encode a signal peptide which allows processing and translocation of the protein, as appropriate, and will usually lack any sequence which might result in the binding of the desired protein of the invention to a membrane. Since, for the most part, the transcriptional initiation region will be for a gene which is expressed and translocated during germination, by employing the signal peptide which provides for translocation, one may also provide for translocation of the protein of interest. In this way, the protein(s) of interest will be translocated from the cells in which they are expressed and may be efficiently harvested. Typically secretion in seeds are across the aleurone or scutellar epithelium layer into the endosperm of the seed. While it is not required that the protein be secreted from the cells in which the protein is produced, this facilitates the isolation and purification of the recombinant protein.

Since the ultimate expression of the desired gene product will be in a eucaryotic cell it is desirable to determine whether any portion of the cloned gene contains sequences which will be processed out as introns by the host's splicosome machinery. If so, site-directed mutagenesis of the "intron" region may be conducted to prevent losing a portion of the genetic message as a false intron code, Reed and Maniatis, *Cell* 41:95-105, 1985.

The vector can be microinjected directly into plant cells by use of micropipettes to mechanically transfer the recombinant DNA. Crossway, *Mol. Gen. Genet*, 202:179-185, 1985. The genetic material may also be transferred into the plant cell by using polyethylene glycol, Krens, et al., *Nature*, 296, 72-74, 1982. Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface, Klein, et al., *Nature*, 327, 70-73, 1987 and Knudsen and Muller, 1991, *Planta*, 185:330-336 teaching particle bombardment of barley endosperm to create transgenic barley. Yet another method of introduction would be fusion of protoplasts with other entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies, Fraley, et al., *Proc. Natl. Acad. Sci. USA*, 79, 1859-1863, 1982.

The vector may also be introduced into the plant cells by electroporation. (Fromm et al., *Proc. Natl. Acad. Sci. USA* 82:5824, 1985). In this technique, plant protoplasts are electroporated in the presence of plasmids containing the gene construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form plant callus.

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum*, and *Datura*.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the heterologous gene is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced from the protoplast suspension.

These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

In some plant cell culture systems, the desired protein of the invention may be excreted or alternatively, the protein may be extracted from the whole plant. Where the desired protein of the invention is secreted into the medium, it may be collected. Alternatively, the embryos and embryoless-half seeds or other plant tissue may be mechanically disrupted to release any secreted protein between cells and tissues. The mixture may be suspended in a buffer solution to retrieve soluble proteins. Conventional protein isolation and purification methods will be then used to purify the recombinant protein. Parameters of time, temperature pH, oxygen, and volumes will be adjusted through routine methods to optimize expression and recovery of heterologous protein.

### iii. Baculovirus Systems

The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector construction employs techniques which are known in the art. Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus-specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media.

After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to

those skilled in the art and fully described in Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987) (hereinafter "Summers and Smith").

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its own set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium. The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT; see Luckow and Summers, *Virology* (1989) 17:31.

The plasmid usually also contains the polyhedrin polyadenylation signal (Miller et al. (1988) *Ann. Rev. Microbiol.*, 42:177) and a prokaryotic ampicillin-resistance (*amp*) gene and origin of replication for selection and propagation in *E. coli*.

Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen et al., (1986) "The Regulation of Baculovirus Gene Expression," in: *The Molecular Biology of Baculoviruses* (ed. Walter Doerfler); EPO

Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlak et al., (1988), *J. Gen. Virol.* 69:765.

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) *Gene*, 73:409). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human (alpha)  $\alpha$ -interferon, Maeda et al., (1985), *Nature* 315:592; human gastrin-releasing peptide, Lebacqz-Verheyden et al., (1988), *Molec. Cell. Biol.* 8:3129; human IL-2, Smith et al., (1985) *Proc. Nat'l Acad. Sci. USA*, 82:8404; mouse IL-3, (Miyajima et al., (1987) *Gene* 58:273; and human glucocerebrosidase, Martin et al. (1988) *DNA*, 7:99, can also be used to provide for secretion in insects.

A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular expression of nonfused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature protein by *in vitro* incubation with cyanogen bromide.

Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5kb section of the baculovirus genome. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art. (See

Summers and Smith *supra*; Ju et al. (1987); Smith et al., *Mol. Cell. Biol.* (1983) 3:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller et al., (1989), *Bioessays* 4:91. The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15  $\mu\text{m}$  in size, are highly refractile, giving them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wild-type virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies. *Current Protocols in Microbiology* Vol. 2 (Ausubel et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, *supra*; Miller et al. (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, *inter alia*: *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni* (PCT Pub. No. WO 89/046699; Carbonell et al., (1985) *J. Virol.* 56:153; Wright (1986) *Nature* 321:718; Smith et al., (1983) *Mol. Cell. Biol.* 3:2156; and see generally, Fraser, et al. (1989) *In Vitro Cell. Dev. Biol.* 25:225).

Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. See, e.g., Summers and Smith *supra*.



The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially free of host debris, e.g., proteins, lipids and polysaccharides.

In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

#### iv. Bacterial Systems

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps

initiate transcription of the lac operon in *Escherichia coli* (*E. coli*) (Raibaud *et al.* (1984) *Annu. Rev. Genet.* 18:173). Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (*lac*) (Chang *et al.* (1977) *Nature* 198:1056), and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (*trp*) (Goeddel *et al.* (1980) *Nuc. Acids Res.* 8:4057; Yelverton *et al.* (1981) *Nucl. Acids Res.* 9:731; U.S. Patent 4,738,921; EPO Publ. Nos. 036 776 and 121 775). The beta-lactamase (*bla*) promoter system (Weissmann (1981) "The cloning of interferon and other mistakes." In *Interferon 3* (ed. I. Gresser)), bacteriophage lambda PL (Shimatake *et al.* (1981) *Nature* 292:128) and T5 (U.S. Patent 4,689,406) promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (U.S. Patent 4,551,433). For example, the *tac* promoter is a hybrid *trp-lac* promoter comprised of both *trp* promoter and *lac* operon sequences that is regulated by the *lac* repressor (Amann *et al.* (1983) *Gene* 25:167; de Boer *et al.* (1983) *Proc. Natl. Acad. Sci.* 80:21). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system (Studier *et al.* (1986) *J. Mol. Biol.* 189:113; Tabor *et al.* (1985) *Proc Natl. Acad. Sci.* 82:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E. coli* operator region (EPO Publ. No. 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon

(Shine *et al.* (1975) *Nature* 254:34). The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' end of *E. coli* 16S rRNA (Steitz *et al.* (1979) "Genetic signals and nucleotide sequences in messenger RNA." In *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberg)). To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site, it is often necessary to optimize the distance between the SD sequence and the ATG of the eukaryotic gene (Sambrook *et al.* (1989) "Expression of cloned genes in *Escherichia coli*." In *Molecular Cloning: A Laboratory Manual*).

A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* or *in vitro* incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site for a processing enzyme (factor Xa) to cleave the bacteriophage protein from the foreign gene (Nagai *et al.* (1984) *Nature* 309:810). Fusion proteins can also be made with sequences from the *lacZ* (Jia *et al.* (1987) *Gene* 60:197), *trpE* (Allen *et al.* (1987) *J. Biotechnol.* 5:93; Makoff *et al.* (1989) *J. Gen. Microbiol.* 135:11), and *Chey* (EPO Publ. No. 324 647) genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated (Miller *et al.* (1989) *Bio/Technology* 7:698).

Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria (U.S. Patent 4,336,336). The

signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either *in vivo* or *in vitro* encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the *E. coli* outer membrane protein gene (*ompA*) (Masui *et al.* (1983), in: *Experimental Manipulation of Gene Expression*; Ghrayeb *et al.* (1984) *EMBO J.* 3:2437) and the *E. coli* alkaline phosphatase signal sequence (*phoA*) (Oka *et al.* (1985) *Proc. Natl. Acad. Sci.* 82:7212). As an additional example, the signal sequence of the alpha-amylase gene from various *Bacillus* strains can be used to secrete heterologous proteins from *B. subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. No. 244 042).

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the *trp* gene in *E. coli* as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number

vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various *Bacillus* strains integrate into the *Bacillus* chromosome (EPO Publ. No. 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline (Davies *et al.* (1978) *Annu. Rev. Microbiol.* 32:469). Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, *inter alia*, the following bacteria: *Bacillus subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541), *Escherichia coli* (Shimatake *et al.* (1981) *Nature* 292:128; Amann *et al.* (1985) *Gene* 40:183; Studier *et al.* (1986) *J. Mol. Biol.* 189:113; EPO Publ. Nos. 036 776, 136 829 and 136 907), *Streptococcus cremoris* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655); *Streptococcus lividans* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655), *Streptomyces lividans* (U.S. Patent 4,745,056).

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with  $\text{CaCl}_2$  or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by

electroporation. Transformation procedures usually vary with the bacterial species to be transformed. (See e.g., use of *Bacillus*: Masson *et al.* (1989) *FEMS Microbiol. Lett.* 60:273; Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541; use of *Campylobacter*: Miller *et al.* (1988) *Proc. Natl. Acad. Sci.* 85:856; and Wang *et al.* (1990) *J. Bacteriol.* 172:949; use of *Escherichia coli*: Cohen *et al.* (1973) *Proc. Natl. Acad. Sci.* 69:2110; Dower *et al.* (1988) *Nucleic Acids Res.* 16:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids. In *Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering* (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) *J. Mol. Biol.* 53:159; Taketo (1988) *Biochim. Biophys. Acta* 949:318; use of *Lactobacillus*: Chassy *et al.* (1987) *FEMS Microbiol. Lett.* 44:173; use of *Pseudomonas*: Fiedler *et al.* (1988) *Anal. Biochem* 170:38; use of *Staphylococcus*: Augustin *et al.* (1990) *FEMS Microbiol. Lett.* 66:203; use of *Streptococcus*: Barany *et al.* (1980) *J. Bacteriol.* 144:698; Harlander (1987) "Transformation of *Streptococcus lactis* by electroporation, in: *Streptococcal Genetics* (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) *Infect. Immun.* 32:1295; Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655; Somkuti *et al.* (1987) *Proc. 4th Evr. Cong. Biotechnology* 1:412.

#### v. Yeast Expression

Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences.

Examples include alcohol dehydrogenase (ADH) (EPO Publ. No. 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO Publ. No. 329 203). The yeast *PHO5* gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:1).

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the *ADH2*, *GAL4*, *GAL10*, OR *PHO5* genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EPO Publ. No. 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, *inter alia*, (Cohen *et al.* (1980) *Proc. Natl. Acad. Sci. USA* 77:1078; Henikoff *et al.* (1981) *Nature* 283:835; Hollenberg *et al.* (1981) *Curr. Topics Microbiol. Immunol.* 96:119; Hollenberg *et al.* (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast *Saccharomyces cerevisiae*," in: *Plasmids of Medical, Environmental and Commercial Importance* (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon *et al.* (1980) *Gene* 11:163; Panthier *et al.* (1980) *Curr. Genet.* 2:109;).

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, plant, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human

superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EPO Publ. No. 196056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (e.g., WO88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EPO Publ. No. 012 873; JPO Publ. No. 62:096,086) and the A-factor gene (U.S. Patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EPO Publ. No. 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (U.S. Patent Nos. 4,546,083 and 4,870,008; EPO Publ. No. 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (See e.g., PCT Publ. No. WO 89/02463.)

Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator



sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein *et al.* (1979) *Gene* 8:17-24), pCl/1 (Brake *et al.* (1984) *Proc. Natl. Acad. Sci USA* 81:4642-4646), and YRp17 (Stinchcomb *et al.* (1982) *J. Mol. Biol.* 158:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Either a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake *et al.*, *supra*.

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver *et al.* (1983) *Methods in Enzymol.* 101:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver *et al.*, *supra*. One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as *ADE2*, *HIS4*, *LEU2*, *TRP1*, and *ALG7*, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of *CUP1* allows yeast to grow in the presence of copper ions (Butt *et al.* (1987) *Microbiol. Rev.* 51:351).

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors and methods of introducing exogenous DNA into yeast hosts have been developed for, *inter alia*, the following yeasts: *Candida albicans* (Kurtz, *et al.* (1986) *Mol. Cell. Biol.* 6:142); *Candida maltosa* (Kunze, *et al.* (1985) *J. Basic Microbiol.* 25:141); *Hansenula polymorpha* (Gleeson, *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302); *Kluyveromyces fragilis* (Das, *et al.* (1984) *J. Bacteriol.* 158:1165); *Kluyveromyces lactis* (De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:737; Van den Berg *et al.* (1990) *Bio/Technology* 8:135); *Pichia guilliermondii* (Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141); *Pichia pastoris* (Cregg, *et al.* (1985) *Mol. Cell. Biol.* 5:3376; U.S. Patent Nos. 4,837,148 and 4,929,555); *Saccharomyces cerevisiae* (Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163); *Schizosaccharomyces pombe* (Beach and Nurse (1981) *Nature* 300:706); and *Yarrowia lipolytica* (Davidow, *et al.* (1985) *Curr. Genet.* 10:380471 Gaillardin, *et al.* (1985) *Curr. Genet.* 10:49).

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., [Kurtz *et al.* (1986) *Mol. Cell. Biol.* 6:142; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; *Candida*]; [Gleeson *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302; *Hansenula*]; [Das *et al.* (1984) *J.*

*Bacteriol.* 158:1165; De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:1165; Van den Berg *et al.* (1990) *Bio/Technology* 8:135; Kluyveromyces]; [Cregg *et al.* (1985) *Mol. Cell. Biol.* 5:3376; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; U.S. Patent Nos. 4,837,148 and 4,929,555; *Pichia*]; [Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163 *Saccharomyces*]; [Beach and Nurse (1981) *Nature* 300:706; *Schizosaccharomyces*]; [Davidow *et al.* (1985) *Curr. Genet.* 10:39; Gaillardin *et al.* (1985) *Curr. Genet.* 10:49; *Yarrowia*].

### Definitions

A composition containing X is "substantially free of" Y when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95% or even 99% by weight.

A "conserved" *Neisseria* amino acid fragment or protein is one that is present in a particular *Neisseria* protein in at least x% of *Neisseria*. The value of x may be 50% or more, e.g., 66%, 75%, 80%, 90%, 95% or even 100% (i.e. the amino acid is found in the protein in question in all *Neisseria*). In order to determine whether an amino acid is "conserved" in a particular *Neisseria* protein, it is necessary to compare that amino acid residue in the sequences of the protein in question from a plurality of different *Neisseria* (a reference population). The reference population may include a number of different *Neisseria* species or may include a single species. The reference population may include a number of different serogroups of a particular species or a single serogroup. A preferred reference population consists of the 5 most common *Neisseria* strains.

The term "heterologous" refers to two biological components that are not found together in nature. The components may be host cells, genes, or regulatory regions, such as promoters. Although the heterologous components are not found together in nature, they can function together, as when a promoter heterologous to a gene is operably linked to the gene. Another example is where a *Neisseria* sequence is heterologous to a mouse host cell.

"Epitope" means antigenic determinant, and may elicit a cellular and/or humoral response.

Conditions for "high stringency" are 65 degrees C in 0.1 xSSC 0.5% SDS solution.

An "origin of replication" is a polynucleotide sequence that initiates and regulates replication of polynucleotides, such as an expression vector. The origin of replication behaves as an autonomous unit of polynucleotide replication within a cell, capable of replication under its own control. An origin of replication may be needed for a vector to replicate in a particular host cell. With certain origins of replication, an expression vector can be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

A "mutant" sequence is defined as a DNA, RNA or amino acid sequence differing from but having homology with the native or disclosed sequence. Depending on the particular sequence, the degree of homology (sequence identity) between the native or disclosed sequence and the mutant sequence is preferably greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) which is calculated as described above. As used herein, an "allelic variant" of a nucleic acid molecule, or region, for which nucleic acid sequence is provided herein is a nucleic acid molecule, or region, that occurs at essentially the same locus in the genome of another or second isolate, and that, due to natural variation caused by, for example, mutation or recombination, has a similar but not identical nucleic acid sequence. A coding region allelic variant typically encodes a protein having similar activity to that of the protein encoded by the gene to which it is being compared. An allelic variant can also comprise an alteration in the 5' or 3' untranslated regions of the gene, such as in regulatory control regions. (see, for example, U.S. Patent 5,753,235).

#### Antibodies

As used herein, the term "antibody" refers to a polypeptide or group of polypeptides composed of at least one antibody combining site. An "antibody combining site" is the three-dimensional binding space with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows a binding of the antibody with the antigen. "Antibody" includes, for example, vertebrate antibodies, hybrid antibodies, chimeric antibodies, humanized antibodies, altered antibodies, univalent antibodies, Fab proteins, and single domain antibodies.

Antibodies against the proteins of the invention are useful for affinity chromatography, immunoassays, and distinguishing/identifying *Neisseria menB* proteins.

Antibodies elicited against the proteins of the present invention bind to antigenic polypeptides or proteins or protein fragments that are present and specifically associated with strains of *Neisseria meningitidis* menB. In some instances, these antigens may be associated with specific strains, such as those antigens specific for the menB strains. The antibodies of the invention may be immobilized to a matrix and utilized in an immunoassay or on an affinity chromatography column, to enable the detection and/or separation of polypeptides, proteins or protein fragments or cells comprising such polypeptides, proteins or protein fragments. Alternatively, such polypeptides, proteins or protein fragments may be immobilized so as to detect antibodies bindably specific thereto.

Antibodies to the proteins of the invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the protein is first used to immunize a suitable animal, preferably a mouse, rat, rabbit or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by *in vitro* immunization using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization. Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000g for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the standard method of Kohler & Milstein (*Nature* (1975) 256:495-96), or a modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate or well coated with the protein antigen. B-cells that express membrane-bound immunoglobulin specific for the antigen bind to the plate, and

are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected MAb-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly  $^{32}\text{P}$  and  $^{125}\text{I}$ ), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct classes, as the same label may serve in several different modes. For example,  $^{125}\text{I}$  may serve as a radioactive label or as an electron-dense reagent. HRP may serve as enzyme or as antigen for a MAb. Further, one may combine various labels for desired effect. For example, MAbs and avidin also require labels in the practice of this invention: thus, one might label a MAb with biotin, and detect its presence with avidin labeled with  $^{125}\text{I}$ , or with an anti-biotin MAb labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

Antigens, immunogens, polypeptides, proteins or protein fragments of the present invention elicit formation of specific binding partner antibodies. These antigens, immunogens, polypeptides, proteins or protein fragments of the present invention comprise immunogenic compositions of the present invention. Such immunogenic compositions may further comprise or include adjuvants, carriers, or other compositions that promote or enhance

or stabilize the antigens, polypeptides, proteins or protein fragments of the present invention. Such adjuvants and carriers will be readily apparent to those of ordinary skill in the art.

### Pharmaceutical Compositions

Pharmaceutical compositions can comprise (include) either polypeptides, antibodies, or nucleic acid of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature, when given to a patient that is febrile. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation can be determined by routine experimentation and is within the judgment of the clinician.

For purposes of the present invention, an effective dose will be from about 0.01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of

organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.

#### Delivery Methods

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal and transcutaneous applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

#### Vaccines

Vaccines according to the invention may either be prophylactic (i.e., to prevent infection) or therapeutic (i.e., to treat disease after infection).

Such vaccines comprise immunizing antigen(s) or immunogen(s), immunogenic polypeptide, protein(s) or protein fragments, or nucleic acids (e.g., ribonucleic acid or deoxyribonucleic acid), usually in combination with "pharmaceutically acceptable carriers," which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or



liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the immunogen or antigen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, etc. pathogens.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT Publ. No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) Rib<sup>TM</sup> adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox<sup>TM</sup>); (3) saponin adjuvants, such as Stimulon<sup>TM</sup> (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (e.g., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.), interferons (e.g., gamma interferon), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., WO 93/13302 and WO 92/19265; and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Alum and MF59 are preferred.

As mentioned above, muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

The vaccine compositions comprising immunogenic compositions (e.g., which may include the antigen, pharmaceutically acceptable carrier, and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Alternatively, vaccine compositions comprising immunogenic compositions may comprise an antigen, polypeptide, protein, protein fragment or nucleic acid in a pharmaceutically acceptable carrier.

More specifically, vaccines comprising immunogenic compositions comprise an immunologically effective amount of the immunogenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, *etc.*), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Typically, the vaccine compositions or immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above under pharmaceutically acceptable carriers.

The immunogenic compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories, and transdermal and transcutaneous applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

As an alternative to protein-based vaccines, DNA vaccination may be employed (e.g., Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648).

### Gene Delivery Vehicles

Gene therapy vehicles for delivery of constructs, including a coding sequence of a therapeutic of the invention, to be delivered to the mammal for expression in the mammal, can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches in *in vivo* or *ex vivo* modality. Expression of such coding sequence can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence *in vivo* can be either constitutive or regulated.

The invention includes gene delivery vehicles capable of expressing the contemplated nucleic acid sequences. The gene delivery vehicle is preferably a viral vector and, more preferably, a retroviral, adenoviral, adeno-associated viral (AAV), herpes viral, or alphavirus vector. The viral vector can also be an astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, or togavirus viral vector. See generally, Jolly (1994) *Cancer Gene Therapy* 1:51-64; Kimura (1994) *Human Gene Therapy* 5:845-852; Connelly (1995) *Human Gene Therapy* 6:185-193; and Kaplitt (1994) *Nature Genetics* 6:148-153.

Retroviral vectors are well known in the art, including B, C and D type retroviruses, xenotropic retroviruses (for example, NZB-X1, NZB-X2 and NZB9-1 (see O'Neill (1985) *J. Virol.* 53:160) polytropic retroviruses e.g., MCF and MCF-MLV (see Kelly (1983) *J. Virol.* 45:291), spumaviruses and lentiviruses. See RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985.

Portions of the retroviral gene therapy vector may be derived from different retroviruses. For example, retrovector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

These recombinant retroviral vectors may be used to generate transduction competent retroviral vector particles by introducing them into appropriate packaging cell lines (see US patent 5,591,624). Retrovirus vectors can be constructed for site-specific integration into host cell DNA by incorporation of a chimeric integrase enzyme into the retroviral particle (see WO96/37626). It is preferable that the recombinant viral vector is a replication defective recombinant virus.

Packaging cell lines suitable for use with the above-described retrovirus vectors are well known in the art, are readily prepared (see WO95/30763 and WO92/05266), and can be used to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles. Preferably, the packaging cell lines are made from human parent cells (e.g., HT1080 cells) or mink parent cell lines, which eliminates inactivation in human serum.

Preferred retroviruses for the construction of retroviral gene therapy vectors include Avian Leukosis Virus, Bovine Leukemia Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis Virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe (1976) *J Virol* 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC No. VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998) and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be obtained from depositories or collections such as the American Type Culture Collection ("ATCC") in Rockville, Maryland or isolated from known sources using commonly available techniques.

Exemplary known retroviral gene therapy vectors employable in this invention include those described in patent applications GB2200651, EP0415731, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5,219,740, US 4,405,712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. See also Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53 (1993) 83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *Proc Natl Acad Sci* 81:6349; and Miller (1990) *Human Gene Therapy* 1.

Human adenoviral gene therapy vectors are also known in the art and employable in this invention. See, for example, Berkner (1988) *Biotechniques* 6:616 and Rosenfeld (1991) *Science* 252:431, and WO93/07283, WO93/06223, and WO93/07282. Exemplary known adenoviral gene therapy vectors employable in this invention include those described in the above referenced documents and in WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102,

WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 and WO95/09654. Alternatively, administration of DNA linked to killed adenovirus as described in Curiel (1992) *Hum. Gene Ther.* 3:147-154 may be employed. The gene delivery vehicles of the invention also include adenovirus associated virus (AAV) vectors. Leading and preferred examples of such vectors for use in this invention are the AAV-2 based vectors disclosed in Srivastava, WO93/09239. Most preferred AAV vectors comprise the two AAV inverted terminal repeats in which the native D-sequences are modified by substitution of nucleotides, such that at least 5 native nucleotides and up to 18 native nucleotides, preferably at least 10 native nucleotides up to 18 native nucleotides, most preferably 10 native nucleotides are retained and the remaining nucleotides of the D-sequence are deleted or replaced with non-native nucleotides. The native D-sequences of the AAV inverted terminal repeats are sequences of 20 consecutive nucleotides in each AAV inverted terminal repeat (i.e., there is one sequence at each end) which are not involved in HP formation. The non-native replacement nucleotide may be any nucleotide other than the nucleotide found in the native D-sequence in the same position. Other employable exemplary AAV vectors are pWP-19, pWN-1, both of which are disclosed in Nahreini (1993) *Gene* 124:257-262. Another example of such an AAV vector is psub201 (see Samulski (1987) *J. Virol.* 61:3096). Another exemplary AAV vector is the Double-D ITR vector. Construction of the Double-D ITR vector is disclosed in US Patent 5,478,745. Still other vectors are those disclosed in Carter US Patent 4,797,368 and Muzyczka US Patent 5,139,941, Chartejee US Patent 5,474,935, and Kotin WO94/288157. Yet a further example of an AAV vector employable in this invention is SSV9AFABTKneo, which contains the AFP enhancer and albumin promoter and directs expression predominantly in the liver. Its structure and construction are disclosed in Su (1996) *Human Gene Therapy* 7:463-470. Additional AAV gene therapy vectors are described in US 5,354,678, US 5,173,414, US 5,139,941, and US 5,252,479.

The gene therapy vectors comprising sequences of the invention also include herpes vectors. Leading and preferred examples are herpes simplex virus vectors containing a sequence encoding a thymidine kinase polypeptide such as those disclosed in US 5,288,641 and EP0176170 (Roizman). Additional exemplary herpes simplex virus vectors include HFEM/ICP6-LacZ disclosed in WO95/04139 (Wistar Institute), pHSVlac described in Geller (1988) *Science* 241:1667-1669 and in WO90/09441 and WO92/07945, HSV Us3::pgC-lacZ

described in Fink (1992) *Human Gene Therapy* 3:11-19 and HSV 7134, 2 RH 105 and GAL4 described in EP 0453242 (Breakefield), and those deposited with the ATCC as accession numbers ATCC VR-977 and ATCC VR-260.

Also contemplated are alpha virus gene therapy vectors that can be employed in this invention. Preferred alpha virus vectors are Sindbis viruses vectors. Togaviruses, Semliki Forest virus (ATCC VR-67; ATCC VR-1247), Middleberg virus (ATCC VR-370), Ross River virus (ATCC VR-373; ATCC VR-1246), Venezuelan equine encephalitis virus (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), and those described in US patents 5,091,309, 5,217,879, and WO92/10578. More particularly, those alpha virus vectors described in U.S. Serial No. 08/405,627, filed March 15, 1995, WO94/21792, WO92/10578, WO95/07994, US 5,091,309 and US 5,217,879 are employable. Such alpha viruses may be obtained from depositories or collections such as the ATCC in Rockville, Maryland or isolated from known sources using commonly available techniques. Preferably, alphavirus vectors with reduced cytotoxicity are used (see USSN 08/679640).

DNA vector systems such as eukaryotic layered expression systems are also useful for expressing the nucleic acids of the invention. See WO95/07994 for a detailed description of eukaryotic layered expression systems. Preferably, the eukaryotic layered expression systems of the invention are derived from alphavirus vectors and most preferably from Sindbis viral vectors.

Other viral vectors suitable for use in the present invention include those derived from poliovirus, for example ATCC VR-58 and those described in Evans, *Nature* 339 (1989) 385 and Sabin (1973) *J. Biol. Standardization* 1:115; rhinovirus, for example ATCC VR-1110 and those described in Arnold (1990) *J Cell Biochem* L401; pox viruses such as canary pox virus or vaccinia virus, for example ATCC VR-111 and ATCC VR-2010 and those described in Fisher-Hoch (1989) *Proc Natl Acad Sci* 86:317; Flexner (1989) *Ann NY Acad Sci* 569:86, Flexner (1990) *Vaccine* 8:17; in US 4,603,112 and US 4,769,330 and WO89/01973; SV40 virus, for example ATCC VR-305 and those described in Mulligan (1979) *Nature* 277:108 and Madzak (1992) *J Gen Virol* 73:1533; influenza virus, for example ATCC VR-797 and recombinant influenza viruses made employing reverse genetics techniques as described in US 5,166,057 and in Enami (1990) *Proc Natl Acad Sci* 87:3802-3805; Enami & Palese (1991) *J Virol* 65:2711-2713 and Luytjes (1989) *Cell* 59:110, (see also McMichael (1983) *NEJ Med* 309:13, and Yap (1978) *Nature* 273:238 and *Nature* (1979) 277:108); human

immunodeficiency virus as described in EP-0386882 and in Buchschacher (1992) *J. Virol.* 66:2731; measles virus, for example ATCC VR-67 and VR-1247 and those described in EP-0440219; Aura virus, for example ATCC VR-368; Bebaru virus, for example ATCC VR-600 and ATCC VR-1240; Cabassou virus, for example ATCC VR-922; Chikungunya virus, for example ATCC VR-64 and ATCC VR-1241; Fort Morgan Virus, for example ATCC VR-924; Getah virus, for example ATCC VR-369 and ATCC VR-1243; Kyzylagach virus, for example ATCC VR-927; Mayaro virus, for example ATCC VR-66; Mucambo virus, for example ATCC VR-580 and ATCC VR-1244; Ndumu virus, for example ATCC VR-371; Pixuna virus, for example ATCC VR-372 and ATCC VR-1245; Tonate virus, for example ATCC VR-925; Trinita virus, for example ATCC VR-469; Una virus, for example ATCC VR-374; Whataroa virus, for example ATCC VR-926; Y-62-33 virus, for example ATCC VR-375; O'Nyong virus, Eastern encephalitis virus, for example ATCC VR-65 and ATCC VR-1242; Western encephalitis virus, for example ATCC VR-70, ATCC VR-1251, ATCC VR-622 and ATCC VR-1252; and coronavirus, for example ATCC VR-740 and those described in Hamre (1966) *Proc Soc Exp Biol Med* 121:190.

Delivery of the compositions of this invention into cells is not limited to the above mentioned viral vectors. Other delivery methods and media may be employed such as, for example, nucleic acid expression vectors, polycationic condensed DNA linked or unlinked to killed adenovirus alone, for example see US Serial No. 08/366,787, filed December 30, 1994 and Curiel (1992) *Hum Gene Ther* 3:147-154 ligand linked DNA, for example see Wu (1989) *J Biol Chem* 264:16985-16987, eucaryotic cell delivery vehicles cells, for example see US Serial No.08/240,030, filed May 9, 1994, and US Serial No. 08/404,796, deposition of photopolymerized hydrogel materials, hand-held gene transfer particle gun, as described in US Patent 5,149,655, ionizing radiation as described in US5,206,152 and in WO92/11033, nucleic charge neutralization or fusion with cell membranes. Additional approaches are described in Philip (1994) *Mol Cell Biol* 14:2411-2418 and in Woffendin (1994) *Proc Natl Acad Sci* 91:1581-1585.

Particle mediated gene transfer may be employed, for example see US Serial No. 60/023,867. Briefly, the sequence can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, as described in Wu &

Wu (1987) *J. Biol. Chem.* 262:4429-4432, insulin as described in Hucked (1990) *Biochem Pharmacol* 40:253-263, galactose as described in Plank (1992) *Bioconjugate Chem* 3:533-539, lactose or transferrin.

Naked DNA may also be employed to transform a host cell. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Uptake efficiency may be improved using biodegradable latex beads. DNA coated latex beads are efficiently transported into cells after endocytosis initiation by the beads. The method may be improved further by treatment of the beads to increase hydrophobicity and thereby facilitate disruption of the endosome and release of the DNA into the cytoplasm.

Liposomes that can act as gene delivery vehicles are described in U.S. 5,422,120, WO95/13796, WO94/23697, WO91/14445 and EP-524,968. As described in USSN. 60/023,867, on non-viral delivery, the nucleic acid sequences encoding a polypeptide can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then be incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, insulin, galactose, lactose, or transferrin. Other delivery systems include the use of liposomes to encapsulate DNA comprising the gene under the control of a variety of tissue-specific or ubiquitously-active promoters. Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin *et al* (1994) *Proc. Natl. Acad. Sci. USA* 91(24):11581-11585. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in U.S. 5,149,655; use of ionizing radiation for activating transferred gene, as described in U.S. 5,206,152 and WO92/11033.

Exemplary liposome and polycationic gene delivery vehicles are those described in US 5,422,120 and 4,762,915; in WO 95/13796; WO94/23697; and WO91/14445; in EP-0524968; and in Stryer, *Biochemistry*, pages 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) *Biochem Biophys Acta* 600:1; Bayer (1979) *Biochem Biophys Acta* 550:464; Rivnay (1987) *Meth Enzymol* 149:119; Wang (1987) *Proc Natl Acad Sci* 84:7851; Plant (1989) *Anal Biochem* 176:420.



A polynucleotide composition can comprises therapeutically effective amount of a gene therapy vehicle, as the term is defined above. For purposes of the present invention, an effective dose will be from about 0.01 mg/ kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

### **Delivery Methods**

Once formulated, the polynucleotide compositions of the invention can be administered (1) directly to the subject; (2) delivered *ex vivo*, to cells derived from the subject; or (3) *in vitro* for expression of recombinant proteins. The subjects to be treated can be mammals or birds. Also, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a tumor or lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in *eg.* WO93/14778. Examples of cells useful in *ex vivo* applications include, for example, stem cells, particularly hematopoietic, lymph cells, macrophages, dendritic cells, or tumor cells.

Generally, delivery of nucleic acids for both *ex vivo* and *in vitro* applications can be accomplished by the following procedures, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

### **Polynucleotide and polypeptide pharmaceutical compositions**

In addition to the pharmaceutically acceptable carriers and salts described above, the following additional agents can be used with polynucleotide and/or polypeptide compositions.

#### A. Polypeptides

One example are polypeptides which include, without limitation: asioloorosomucoid (ASOR); transferrin; asialoglycoproteins; antibodies; antibody fragments; ferritin; interleukins; interferons, granulocyte, macrophage colony stimulating factor (GM-CSF),

granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor and erythropoietin. Viral antigens, such as envelope proteins, can also be used. Also, proteins from other invasive organisms, such as the 17 amino acid peptide from the circumsporozoite protein of plasmodium falciparum known as RII.

#### B.Hormones, Vitamins, Etc.

Other groups that can be included are, for example: hormones, steroids, androgens, estrogens, thyroid hormone, or vitamins, folic acid.

#### C.Polyalkylenes, Polysaccharides, etc.

Also, polyalkylene glycol can be included with the desired polynucleotides or polypeptides. In a preferred embodiment, the polyalkylene glycol is polyethylene glycol. In addition, mono-, di-, or polysaccharides can be included. In a preferred embodiment of this aspect, the polysaccharide is dextran or DEAE-dextran. Also, chitosan and poly(lactide-co-glycolide)

#### D.Lipids, and Liposomes

The desired polynucleotide or polypeptide can also be encapsulated in lipids or packaged in liposomes prior to delivery to the subject or to cells derived therefrom.

Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed polynucleotide or polypeptide to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight (1991) *Biochim. Biophys. Acta.* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (Malone (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); and purified transcription factors (Debs (1990) *J. Biol. Chem.* 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner *supra*). Other commercially available liposomes include transfectace (DDAB/DOPE) and

DOTAP/DOPE (Boehringer). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, eg. Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; WO90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

The liposomes can comprise multilammellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See eg. Straubinger (1983) *Meth. Immunol.* 101:512-527; Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; Papahadjopoulos (1975) *Biochim. Biophys. Acta* 394:483; Wilson (1979) *Cell* 17:77; Deamer & Bangham (1976) *Biochim. Biophys. Acta* 443:629; Ostro (1977) *Biochem. Biophys. Res. Commun.* 76:836; Fraley (1979) *Proc. Natl. Acad. Sci. USA* 76:3348; Enoch & Strittmatter (1979) *Proc. Natl. Acad. Sci. USA* 76:145; Fraley (1980) *J. Biol. Chem.* (1980) 255:10431; Szoka & Papahadjopoulos (1978) *Proc. Natl. Acad. Sci. USA* 75:145; and Schaefer-Ridder (1982) *Science* 215:166.

#### E.Lipoproteins

In addition, lipoproteins can be included with the polynucleotide or polypeptide to be delivered. Examples of lipoproteins to be utilized include: chylomicrons, HDL, IDL, LDL, and VLDL. Mutants, fragments, or fusions of these proteins can also be used. Also, modifications of naturally occurring lipoproteins can be used, such as acetylated LDL. These lipoproteins can target the delivery of polynucleotides to cells expressing lipoprotein receptors. Preferably, if lipoproteins are including with the polynucleotide to be delivered, no other targeting ligand is included in the composition.

Naturally occurring lipoproteins comprise a lipid and a protein portion. The protein portion are known as apoproteins. At the present, apoproteins A, B, C, D, and E have been

isolated and identified. At least two of these contain several proteins, designated by Roman numerals, AI, AII, AIV; CI, CII, CIII.

A lipoprotein can comprise more than one apoprotein. For example, naturally occurring chylomicrons comprises of A, B, C, and E, over time these lipoproteins lose A and acquire C and E apoproteins. VLDL comprises A, B, C, and E apoproteins, LDL comprises apoprotein B; and HDL comprises apoproteins A, C, and E.

The amino acid of these apoproteins are known and are described in, for example, Breslow (1985) *Annu Rev. Biochem* 54:699; Law (1986) *Adv. Exp Med. Biol.* 151:162; Chen (1986) *J Biol Chem* 261:12918; Kane (1980) *Proc Natl Acad Sci USA* 77:2465; and Utermann (1984) *Hum Genet* 65:232.

Lipoproteins contain a variety of lipids including, triglycerides, cholesterol (free and esters), and phospholipids. The composition of the lipids varies in naturally occurring lipoproteins. For example, chylomicrons comprise mainly triglycerides. A more detailed description of the lipid content of naturally occurring lipoproteins can be found, for example, in *Meth. Enzymol.* 128 (1986). The composition of the lipids are chosen to aid in conformation of the apoprotein for receptor binding activity. The composition of lipids can also be chosen to facilitate hydrophobic interaction and association with the polynucleotide binding molecule.

Naturally occurring lipoproteins can be isolated from serum by ultracentrifugation, for instance. Such methods are described in *Meth. Enzymol. (supra)*; Pitas (1980) *J. Biochem.* 255:5454-5460 and Mahey (1979) *J Clin. Invest* 64:743-750.

Lipoproteins can also be produced by *in vitro* or recombinant methods by expression of the apoprotein genes in a desired host cell. See, for example, Atkinson (1986) *Annu Rev Biophys Chem* 15:403 and Radding (1958) *Biochim Biophys Acta* 30: 443.

Lipoproteins can also be purchased from commercial suppliers, such as Biomedical Technologies, Inc., Stoughton, Massachusetts, USA.

Further description of lipoproteins can be found in Zuckermann et al., PCT. Appln. No. US97/14465.

#### F.Polycationic Agents

Polycationic agents can be included, with or without lipoprotein, in a composition with the desired polynucleotide or polypeptide to be delivered.

Polycationic agents, typically, exhibit a net positive charge at physiological relevant pH and are capable of neutralizing the electrical charge of nucleic acids to facilitate delivery to a desired location. These agents have both in vitro, ex vivo, and in vivo applications. Polycationic agents can be used to deliver nucleic acids to a living subject either intramuscularly, subcutaneously, etc.

The following are examples of useful polypeptides as polycationic agents: polylysine, polyarginine, polyornithine, and protamine. Other examples include histones, protamines, human serum albumin, DNA binding proteins, non-histone chromosomal proteins, coat proteins from DNA viruses, such as (X174, transcriptional factors also contain domains that bind DNA and therefore may be useful as nucleic acid condensing agents. Briefly, transcriptional factors such as C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP, and TFIID contain basic domains that bind DNA sequences.

Organic polycationic agents include: spermine, spermidine, and putrescine.

The dimensions and of the physical properties of a polycationic agent can be extrapolated from the list above, to construct other polypeptide polycationic agents or to produce synthetic polycationic agents.

#### Synthetic Polycationic Agents

Synthetic polycationic agents which are useful include, for example, DEAE-dextran, polybrene. Lipofectin<sup>®</sup>, and lipofectAMINE<sup>®</sup> are monomers that form polycationic complexes when combined with polynucleotides or polypeptides.

#### Immunodiagnostic Assays

Neisserial antigens of the invention can be used in immunoassays to detect antibody levels (or, conversely, anti-Neisserial antibodies can be used to detect antigen levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to Neisserial proteins within biological samples, including for example, blood or serum samples, can be detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known; examples of which are

assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, *etc.*) required for the conduct of the assay, as well as suitable set of assay instructions.

### Nucleic Acid Hybridisation

"Hybridization" refers to the association of two nucleic acid sequences to one another by hydrogen bonding. Typically, one sequence will be fixed to a solid support and the other will be free in solution. Then, the two sequences will be placed in contact with one another under conditions that favor hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase sequence to the solid support (Denhardt's reagent or BLOTTO); concentration of the sequences; use of compounds to increase the rate of association of sequences (dextran sulfate or polyethylene glycol); and the stringency of the washing conditions following hybridization. See Sambrook *et al.* [*supra*] Volume 2, chapter 9, pages 9.47 to 9.57.

"Stringency" refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 120 to 200°C below the calculated  $T_m$  of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook *et al.* at page 9.50.

Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1 µg for a plasmid or phage digest to  $10^{-9}$  to  $10^{-8}$  g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides, substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a single-copy

yeast gene can be detected with an exposure time of only 1 hour starting with 1  $\mu$ g of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of  $10^8$  cpm/ $\mu$ g. For a single-copy mammalian gene a conservative approach would start with 10  $\mu$ g of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than  $10^8$  cpm/ $\mu$ g, resulting in an exposure time of ~24 hours.

Several factors can affect the melting temperature ( $T_m$ ) of a DNA-DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

$$T_m = 81 + 16.6(\log_{10} C_i) + 0.4[\%(G + C)] - 0.6(\% \text{formamide}) - 600/n - 1.5(\% \text{mismatch}).$$

where  $C_i$  is the salt concentration (monovalent ions) and  $n$  is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) *Anal. Biochem.* 138: 267-284).

In designing a hybridization experiment, some factors affecting nucleic acid hybridization can be conveniently altered. The temperature of the hybridization and washes and the salt concentration during the washes are the simplest to adjust. As the temperature of the hybridization increases (*ie.* stringency), it becomes less likely for hybridization to occur between strands that are nonhomologous, and as a result, background decreases. If the radiolabeled probe is not completely homologous with the immobilized fragment (as is frequently the case in gene family and interspecies hybridization experiments), the hybridization temperature must be reduced, and background will increase. The temperature of the washes affects the intensity of the hybridizing band and the degree of background in a similar manner. The stringency of the washes is also increased with decreasing salt concentrations.

In general, convenient hybridization temperatures in the presence of 50% formamide are 42°C for a probe with is 95% to 100% homologous to the target fragment, 37°C for 90% to 95% homology, and 32°C for 85% to 90% homology. For lower homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be

washed at high stringency and reexposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

### **Nucleic Acid Probe Assays**

Methods such as PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes according to the invention can determine the presence of cDNA or mRNA. A probe is said to "hybridize" with a sequence of the invention if it can form a duplex or double stranded complex, which is stable enough to be detected.

The nucleic acid probes will hybridize to the Neisserial nucleotide sequences of the invention (including both sense and antisense strands). Though many different nucleotide sequences will encode the amino acid sequence, the native Neisserial sequence is preferred because it is the actual sequence present in cells. mRNA represents a coding sequence and so a probe should be complementary to the coding sequence; single-stranded cDNA is complementary to mRNA, and so a cDNA probe should be complementary to the non-coding sequence.

The probe sequence need not be identical to the Neisserial sequence (or its complement) — some variation in the sequence and length can lead to increased assay sensitivity if the nucleic acid probe can form a duplex with target nucleotides, which can be detected. Also, the nucleic acid probe can include additional nucleotides to stabilize the formed duplex. Additional Neisserial sequence may also be helpful as a label to detect the formed duplex. For example, a non-complementary nucleotide sequence may be attached to the 5' end of the probe, with the remainder of the probe sequence being complementary to a Neisserial sequence. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the a Neisserial sequence in order to hybridize therewith and thereby form a duplex which can be detected.

The exact length and sequence of the probe will depend on the hybridization conditions, such as temperature, salt condition and the like. For example, for diagnostic applications, depending on the complexity of the analyte sequence, the nucleic acid probe typically contains at least 10-20 nucleotides, preferably 15-25, and more preferably at least 30 nucleotides, although it may be shorter than this. Short primers generally require cooler temperatures to form sufficiently stable hybrid complexes with the template.



Probes may be produced by synthetic procedures, such as the triester method of Matteucci *et al.* [*J. Am. Chem. Soc.* (1981) 103:3185], or according to Urdea *et al.* [*Proc. Natl. Acad. Sci. USA* (1983) 80: 7461], or using commercially available automated oligonucleotide synthesizers.

The chemical nature of the probe can be selected according to preference. For certain applications, DNA or RNA are appropriate. For other applications, modifications may be incorporated *eg.* backbone modifications, such as phosphorothioates or methylphosphonates, can be used to increase *in vivo* half-life, alter RNA affinity, increase nuclease resistance *etc.* [*eg.* see Agrawal & Iyer (1995) *Curr Opin Biotechnol* 6:12-19; Agrawal (1996) *TIBTECH* 14:376-387]; analogues such as peptide nucleic acids may also be used [*eg.* see Corey (1997) *TIBTECH* 15:224-229; Buchardt *et al.* (1993) *TIBTECH* 11:384-386].

One example of a nucleotide hybridization assay is described by Urdea *et al.* in international patent application WO92/02526 [see also US patent 5,124,246].

Alternatively, the polymerase chain reaction (PCR) is another well-known means for detecting small amounts of target nucleic acids. The assay is described in: Mullis *et al.* [*Meth. Enzymol.* (1987) 155: 335-350]; US patent 4,683,195; and US patent 4,683,202. Two "primer" nucleotides hybridize with the target nucleic acids and are used to prime the reaction. The primers can comprise sequence that does not hybridize to the sequence of the amplification target (or its complement) to aid with duplex stability or, for example, to incorporate a convenient restriction site. Typically, such sequence will flank the desired Neisserial sequence.

A thermostable polymerase creates copies of target nucleic acids from the primers using the original target nucleic acids as a template. After a threshold amount of target nucleic acids are generated by the polymerase, they can be detected by more traditional methods, such as Southern blots. When using the Southern blot method, the labelled probe will hybridize to the Neisserial sequence (or its complement).

Also, mRNA or cDNA can be detected by traditional blotting techniques described in Sambrook *et al* [*supra*]. mRNA, or cDNA generated from mRNA using a polymerase enzyme, can be purified and separated using gel electrophoresis. The nucleic acids on the gel are then blotted onto a solid support, such as nitrocellulose. The solid support is exposed to a labelled probe and then washed to remove any unhybridized probe. Next, the duplexes

containing the labeled probe are detected. Typically, the probe is labelled with a radioactive moiety.

## EXAMPLES

The examples describe nucleic acid sequences which have been identified in *N. meningitidis*, and *N. gonorrhoeae* along with their respective and putative translation products. Not all of the nucleic acid sequences are complete *ie.* they encode less than the full-length wild-type protein.

The examples are generally in the following format:

- a nucleotide sequence which has been identified in *N. meningitidis*
- the putative translation product of said *N. meningitidis* sequence
- a computer analysis of said translation product based on database comparisons
- a corresponding nucleotide sequence identified from *N. gonorrhoeae*
- the putative translation product of said *N. gonorrhoeae* sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a corresponding nucleotide sequence identified from strain A of *N. meningitidis*
- the putative translation product of said *N. meningitidis* strain A sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a description of the characteristics of the protein which indicates that it might be suitably antigenic or immunogenic.

Sequence comparisons were performed at NCBI (<http://www.ncbi.nlm.nih.gov>) using the algorithms BLAST, BLAST2, BLASTn, BLASTp, tBLASTn, BLASTx, & tBLASTx [eg. see also Altschul *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:2289-3402]. Searches were performed against the following databases: non-redundant GenBank+EMBL+DDBJ+PDB sequences and non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR sequences.

Dots within nucleotide sequences represent nucleotides which have been arbitrarily introduced in order to maintain a reading frame. In the same way, double-underlined nucleotides were removed. Lower case letters represent ambiguities which arose during

alignment of independent sequencing reactions (some of the nucleotide sequences in the examples are derived from combining the results of two or more experiments).

Nucleotide sequences were scanned in all six reading frames to predict the presence of hydrophobic domains using an algorithm based on the statistical studies of Esposti *et al.* [Critical evaluation of the hydropathy of membrane proteins (1990) *Eur J Biochem* 190:207-219]. These domains represent potential transmembrane regions or hydrophobic leader sequences.

Open reading frames were predicted from fragmented nucleotide sequences using the program ORFFINDER (NCBI).

Underlined amino acid sequences indicate possible transmembrane domains or leader sequences in the ORFs, as predicted by the PSORT algorithm (<http://www.psort.nibb.ac.jp>). Functional domains were also predicted using the MOTIFS program (GCG Wisconsin & PROSITE).

For each of the following examples: based on the presence of a putative leader sequence and/or several putative transmembrane domains (single-underlined) in the gonococcal protein, it is predicted that the proteins from *N. meningitidis* and *N. gonorrhoeae*, and their respective epitopes, could be useful antigens or immunogenic compositions for vaccines or diagnostics.

The standard techniques and procedures which may be employed in order to perform the invention (e.g. to utilize the disclosed sequences for vaccination or diagnostic purposes) were summarized above. This summary is not a limitation on the invention but, rather, gives examples that may be used, but are not required.

In particular, the following methods were used to express, purify and biochemically characterize the proteins of the invention.

#### **Chromosomal DNA Preparation**

*N.meningitidis* strain 2996 was grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20%(w/v) Sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one CHCl<sub>3</sub>/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation.

The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading the OD at 260 nm.

### Oligonucleotide design

Synthetic oligonucleotide primers were designed on the basis of the coding sequence of each ORF, using (a) the meningococcus B sequence when available, or (b) the gonococcus/meningococcus A sequence, adapted to the codon preference usage of meningococcus as necessary. Any predicted signal peptides were omitted, by designing the 5' primers to sequence immediately downstream from the predicted leader sequence.

For most ORFs, the 5' primers included two restriction enzyme recognition sites (*Bam*HI-*Nde*I, *Bam*HI-*Nhe*I, *Eco*RI-*Nde*I or *Eco*RI-*Nhe*I), depending on the restriction pattern of the gene of interest. The 3' primers included a *Xho*I or a *Hind*III restriction site (table 1). This procedure was established in order to direct the cloning of each amplification product (corresponding to each ORF) into two different expression systems: pGEX-KG (using *Bam*HI-*Xho*I, *Bam*HI-*Hind*III, *Eco*RI-*Xho*I or *Eco*RI-*Hind*III), and pET21b+ (using *Nde*I-*Xho*I, *Nhe*I-*Xho*I, *Nde*I-*Hind*III or *Nhe*I-*Hind*III).

5'-end primer tail:	<u>CGCGGATCCCATATG</u>	( <i>Bam</i> HI- <i>Nde</i> I )
	<u>CGCGGATCCGCTAGC</u>	( <i>Bam</i> HI- <i>Nhe</i> I)
	<u>CCGGAATTCTACATATG</u>	( <i>Eco</i> RI- <i>Nde</i> I)
	<u>CCGGAATTCTAGCTAGC</u>	( <i>Eco</i> RI- <i>Nhe</i> I)
3'-end primer tail:	<u>CCCGCTCGAG</u>	( <i>Xho</i> I)
	<u>CCCGCTCGAG</u>	( <i>Hind</i> III)

For cloning ORFs into the pGEX-His vector, the 5' and 3' primers contained only one restriction enzyme site (*Eco*RI, *Kpn*I or *Sal*I for the 5' primers and *Pst*I, *Xba*I, *Sph*I or *Sal*I for the 3' primers). Again restriction sites were chosen according to the particular restriction pattern of the gene (table 1).

5'-end primer tail:	(AAA) AAAGAATTC	( <i>Eco</i> RI )
	(AAA) AAAGGTACC	( <i>Kpn</i> I)
3'-end primer tail:	(AAA) AAATGTCAG	( <i>Pst</i> I)
	(AAA) AAATCTAGA	( <i>Xba</i> I)

5' or 3'-end primer tail: AAAGCATGC

(*Sph*I)  
AAAAAAGTCGAC (*Sal*II)

As well as containing the restriction enzyme recognition sequences, the primers included nucleotides which hybridized to the sequence to be amplified. The melting temperature depended on the number and type of hybridising nucleotides in the whole primer, and was determined for each primer using the formulae:

$$T_m = 4 (G+C) + 2 (A+T) \quad (\text{tail excluded})$$

$$T_m = 64.9 + 0.41 (\% \text{ GC}) - 600/N \quad (\text{whole primer})$$

The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-55°C for the hybridising region alone.

Table 1 shows the forward and reverse primers used for each amplification. In certain cases, the sequence of the primer does not exactly match the sequence of the predicted ORF. This is because when initial amplifications were performed, the complete 5' and/or 3' sequences for some meningococcal B ORFs were not known. However the corresponding sequences had been identified in *Gonococcus* or in *Meningococcus A*. Hence, when the *Meningococcus B* sequence was incomplete or uncertain, *Gonococcal* or *Meningococcal A* sequences were used as the basis for primer design. These sequences were altered to take account of codon preference. It can be appreciated that, once the complete sequence is identified, this approach will no longer be necessary.

Oligonucleotides were synthesized using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH<sub>4</sub>OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in either 100µl or 1.0ml of water. The OD<sub>260</sub> was determined using a Perkin Elmer Lambda Bio spectrophotometer and the concentration adjusted to 2-10pmol/µl.

### Amplification

The standard PCR protocol was as follows: 50-200ng of genomic DNA was used as a template in the presence of 20-40µM of each oligonucleotide primer, 400-800µM dNTPs solution, 1x PCR buffer (including 1.5mM MgCl<sub>2</sub>), 2.5 units *TaqI* DNA polymerase (using

Perkin-Elmer AmpliTaq, GIBCO Platinum, Pwo DNA polymerase, or Tahara Shuzo Taq polymerase). In some cases, PCR was optimised by the addition of 10µl DMSO or 50µl 2M Betaine.

After a hot start (adding the polymerase during a preliminary 3 minute incubation of the whole mix at 95°C), each sample underwent a two-step amplification. The first 5 cycles were performed using the hybridization temperature that excluded the restriction enzyme tail of the primer (see above). This was followed by 30 cycles using the hybridization temperature calculated for the whole length oligos. The cycles were completed with a 10 minute extension step at 72°C. The standard cycles were as follows:

	Denaturation	Hybridisation	Elongation
First 5 cycles	30 seconds 95°C	30 seconds 50-55°C	30-60 seconds 72°C
Last 30 cycles	30 seconds 95°C	30 seconds 65-70°C	30-60 seconds 72°C

Elongation times varied according to the length of the ORF to be amplified. Amplifications were performed using either a 9600 or a 2400 Perkin Elmer GeneAmp PCR System. To check the results, 1/10 of the amplification volume was loaded onto a 1-1.5% (w/v) agarose gel and the size of each amplified fragment compared with a DNA molecular weight marker.

The amplified DNA was either loaded directly on a 1% agarose gel or first precipitated with ethanol and resuspended in a volume suitable to be loaded on a 1.0% agarose gel. The DNA fragment corresponding to the band of correct size was purified using the Qiagen Gel Extraction Kit, following the manufacturer's protocol. DNA fragments were eluted in a volume of 30µl or 50µl with either H<sub>2</sub>O or 10mM Tris, pH 8.5.

#### **Digestion of PCR fragments**

The purified DNA corresponding to the amplified fragment was doubly-digested with the appropriate restriction enzymes for; cloning into pET-21b+ and expressing the protein as a C-terminus His-tagged fusion, for cloning into pGEX-KG and expressing the protein as a N-

terminus GST-fusion, and for cloning into pGEX-His and expressing the protein as a N-terminus GST-His tagged fusion.

Each purified DNA fragment was incubated at 37°C for 3 hours to overnight with 20 units of appropriate restriction enzyme (New England Biolabs) in a volume of either 30 or 40µl in the presence of suitable digestion buffer. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted in a volume of 30µl or 50µl with either H<sub>2</sub>O or 10mM Tris, pH 8.5. The DNA concentration was determined by quantitative agarose gel electrophoresis (1.0% gel) in the presence of a titrated molecular weight marker.

#### **Digestion of the cloning vectors (pET22B, pGEX-KG, pTRC-His A, pET21b+, pGEX-KG, and pGEX-His)**

The vector pGEX-His is a modified pGEX-2T vector carrying a region encoding six histidine residues upstream of the thrombin cleavage site and containing the multiple cloning site of the vector pTRC99 (Pharmacia). 10 µg plasmid was double-digested with 50 units of each restriction enzyme in 200 µl reaction volume in the presence of appropriate buffer by overnight incubation at 37°C. After loading the whole digestion on a 1% agarose gel, the band corresponding to the digested vector was purified from the gel using the Qiagen QIAquick Gel Extraction Kit and the DNA was eluted in 50 µl of 10 mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD<sub>260</sub> of the sample, and adjusted to 50 µg/µl. 1 µl of plasmid was used for each cloning procedure.

10µg of plasmid vector was doubly-digested with 50 units of each restriction enzyme in a volume of 200µl with the appropriate buffer overnight at 37°C. The digest was loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit. DNA was eluted in 50µl of 10mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD<sub>260nm</sub> and the concentration adjusted to 50µg/µl. 1µl of plasmid was used for each cloning procedure.

#### **Cloning**

For some ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated in both pET22b and pGEX-KG. In a final volume of 20 µl, a molar

ratio of 3:1 fragment/vector was ligated using 0.5  $\mu$ l of NEB T4 DNA ligase (400 units/ $\mu$ l), in the presence of the buffer supplied by the manufacturer. The reaction was incubated at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit", following the manufacturer's instructions.

In order to introduce the recombinant plasmid in a suitable strain, 100  $\mu$ l *E. coli* DH5 competent cells were incubated with the ligase reaction solution for 40 minutes on ice, then at 37°C for 3 minutes, then, after adding 800  $\mu$ l LB broth, again at 37°C for 20 minutes. The cells were then centrifuged at maximum speed in an Eppendorf microfuge and resuspended in approximately 200  $\mu$ l of the supernatant. The suspension was then plated on LB ampicillin (100 mg/ml ).

The screening of the recombinant clones was performed by growing 5 randomly-chosen colonies overnight at 37 °C in either 2 ml (pGEX or pTC clones) or 5ml (pET clones) LB broth + 100  $\mu$ g/ml ampicillin. The cells were then pelleted and the DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions, to a final volume of 30  $\mu$ l. 5  $\mu$ l of each individual miniprep (approximately 1g ) were digested with either *NdeI/XhoI* or *BamHI/XhoI* and the whole digestion loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1Kb DNA Ladder, GIBCO). The screening of the positive clones was made on the base of the correct insert size.

For other ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated into both pET21b+ and pGEX-KG. A molar ratio of of 3:1 fragment/vector was used in a final volume of 20 $\mu$ l, that included 0.5 $\mu$ l T4 DNA ligase (400 units/ $\mu$ l, NEB) and ligation buffer supplied by the manufacturer. The reaction was performed at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit" and the manufacturer's protocol.

Recombinant plasmid was transformed into 100 $\mu$ l of competent *E. coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice then at 37°C for 3 minutes. This was followed by the addition of 800 $\mu$ l LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200 $\mu$ l of the supernatant and plated onto LB ampicillin (100mg/ml ) agar.



Screening for recombinant clones was performed by growing 5 randomly selected colonies overnight at 37°C in either 2.0ml (pGEX-KG clones) or 5.0ml (pET clones) LB broth + 100µg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1µg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

ORFs were cloned into PGEX-His, by doubly-digesting the PCR product and ligating into similarly digested vector. After cloning, recombinant plasmids were transformed into the *E.coli* host W3110. Individual clones were grown overnight at 37°C in LB broth with 50µg/ml ampicillin.

Certain ORFs may be cloned into the pGEX-HIS vector using *EcoRI-PstI* cloning sites, or *EcoRI-SalI*, or *SalI-PstI*. After cloning, the recombinant plasmids may be introduced in the *E.coli* host W3110.

### Expression

Each ORF cloned into the expression vector may then be transformed into the strain suitable for expression of the recombinant protein product. 1 µl of each construct was used to transform 30 µl of *E.coli* BL21 (pGEX vector), *E.coli* TOP 10 (pTRC vector) or *E.coli* BL21-DE3 (pET vector), as described above. In the case of the pGEX-His vector, the same *E.coli* strain (W3110) was used for initial cloning and expression. Single recombinant colonies were inoculated into 2ml LB+Amp (100 µg/ml), incubated at 37°C overnight, then diluted 1:30 in 20 ml of LB+Amp (100 µg/ml) in 100 ml flasks, making sure that the OD<sub>600</sub> ranged between 0.1 and 0.15. The flasks were incubated at 30°C into gyratory water bath shakers until OD indicated exponential growth suitable for induction of expression (0.4-0.8 OD for pET and pTRC vectors; 0.8-1 OD for pGEX and pGEX-His vectors). For the pET, pTRC and pGEX-His vectors, the protein expression was induced by addition of 1mM IPTG, whereas in the case of pGEX system the final concentration of IPTG was 0.2 mM. After 3 hours incubation at 30°C, the final concentration of the sample was checked by OD. In order to check expression, 1ml of each sample was removed, centrifuged in a microfuge, the pellet

resuspended in PBS, and analysed by 12% SDS-PAGE with Coomassie Blue staining. The whole sample was centrifuged at 6000g and the pellet resuspended in PBS for further use.

#### **GST-fusion proteins large-scale purification.**

For some ORFs, a single colony was grown overnight at 37°C on LB+Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600 ml of fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD<sub>550</sub> 0.8-1. Protein expression was induced with 0.2mM IPTG followed by three hours incubation. The culture was centrifuged at 8000 rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml cold PBS. The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia) (previously washed with PBS) and incubated at room temperature for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold PBS for 10 minutes, resuspended in 1ml cold PBS, and loaded on a disposable column. The resin was washed twice with 2ml cold PBS until the flow-through reached OD<sub>280</sub> of 0.02-0.06. The GST-fusion protein was eluted by addition of 700µl cold Glutathione elution buffer 10mM reduced glutathione, 50mM Tris-HCl) and fractions collected until the OD<sub>280</sub> was 0.1. 21µl of each fraction were loaded on a 12% SDS gel using either Biorad SDS-PAGE Molecular weight standard broad range (M1) (200, 116.25, 97.4, 66.2, 45, 31, 21.5, 14.4, 6.5 kDa) or Amersham Rainbow Marker (M'') (220, 66, 46, 30, 21.5, 14.3 kDa) as standards. As the MW of GST is 26kDa, this value must be added to the MW of each GST-fusion protein.

For other ORFs, for each clone to be purified as a GST-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp. (100µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD<sub>550nm</sub> reached 0.6-0.8. Recombinant protein expression was induced by addition of IPTG (final concentration 0.2mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml cold PBS. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia), previously equilibrated with PBS, and incubated at room temperature with gentle agitation for 30 min. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batchwise) with 10ml cold PBS for 10 min, resuspended in 1ml cold PBS, and loaded onto a disposable column. The resin continued to be washed with cold PBS, until the OD<sub>280nm</sub> of the flow-through reached 0.02-0.01. The GST-fusion protein was eluted by addition of 700µl cold glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl pH 8.0) and fractions collected, until the OD<sub>280nm</sub> of the eluate indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. The molecular mass of the purified proteins was determined using either the Bio-Rad broad range molecular weight standard (M1) (200, 116, 97.4, 66.2, 45.0, 31.0, 21.5, 14.4, 6.5 kDa) or the Amersham Rainbow Marker (M2) (220, 66.2, 46.0, 30.0, 21.5, 14.3 kDa). The molecular weights of GST-fusion proteins are a combination of the 26 kDa GST protein and its fusion partner. Protein concentrations were estimated using the Bradford assay.

#### **His-fusion soluble proteins large-scale purification.**

For some ORFs, a single colony was grown overnight at 37°C on a LB + Amp agar plate. The bacteria were inoculated into 20ml of LB+Amp liquid culture and incubated overnight in a water bath shaker. Bacteria were diluted 1:30 into 600ml fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD<sub>550</sub> 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000 rpm at 4°C, the supernatant was discarded and the bacterial pellet was resuspended in 7.5ml cold 10mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Ni<sup>2+</sup>-resin (Pharmacia) (previously washed with 10mM imidazole buffer) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold 10mM imidazole buffer for 10 minutes,

resuspended in 1ml cold 10mM imidazole buffer and loaded on a disposable column. The resin was washed at 4°C with 2ml cold 10mM imidazole buffer until the flow-through reached the O.D<sub>280</sub> of 0.02-0.06. The resin was washed with 2ml cold 20mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 20 mM imidazole, pH 8) until the flow-through reached the O.D<sub>280</sub> of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl cold 250mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 250 mM imidazole, pH 8) and fractions collected until the O.D<sub>280</sub> was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

#### **His-fusion insoluble proteins large-scale purification.**

A single colony was grown overnight at 37 °C on a LB + Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600ml fresh medium and let to grow at the optimal temperature (37°C) to O.D<sub>550</sub> 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml buffer B (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed twice and centrifuged again. The supernatant was stored at -20°C, while the pellets were resuspended in 2 ml guanidine buffer (6M guanidine hydrochloride, 100mM phosphate buffer, 10 mM Tris-HCl, pH 7.5) and treated in a homogenizer for 10 cycles. The product was centrifuged at 13000 rpm for 40 minutes. The supernatant was mixed with 150µl Ni<sup>2+</sup>-resin (Pharmacia) (previously washed with buffer B) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700 g for 5 minutes at 4°C. The resin was washed twice with 10 ml buffer B for 10 minutes, resuspended in 1ml buffer B, and loaded on a disposable column. The resin was washed at room temperature with 2ml buffer B until the flow-through reached the OD<sub>280</sub> of 0.02-0.06. The resin was washed with 2ml buffer C (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the flow-through reached the O.D<sub>280</sub> of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl elution buffer (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions collected until the OD<sub>280</sub> was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

**Purification of His-fusion proteins.**

For each clone to be purified as a His-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp (100 µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100 µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD<sub>550nm</sub> reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml of either (i) cold buffer A (300mM NaCl, 50mM phosphate buffer, 10mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8) for insoluble proteins. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6M guanidine hydrochloride, 100mM phosphate buffer, 10mM Tris-HCl, pH 7.5) and treated with a Dounce homogenizer for 10 cycles. The homogenate was centrifuged at 13 000xg for 40 min and the supernatant retained.

Supernatants for both soluble and insoluble preparations were mixed with 150µl Ni<sup>2+</sup>-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to manufacturers protocol. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD<sub>280nm</sub> of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the the OD<sub>280nm</sub> of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700µl of either (i) cold elution buffer A (300mM NaCl, 50mM phosphate buffer, 250mM imidazole, pH 8.0) or (ii) elution buffer B (8 M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions

collected until the  $O.D_{280nm}$  indicated all the recombinant protein was obtained. 20 $\mu$ l aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. Protein concentrations were estimated using the Bradford assay.

### His-fusion proteins renaturation

In the cases where denaturation was required to solubilize proteins, a renaturation step was employed prior to immunization. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10%(v/v). The proteins were diluted to 200 $\mu$ g/ml using dialysis buffer I (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Alternatively, 10% glycerol was added to the denatured proteins. The proteins were then diluted to 20 $\mu$ g/ml using dialysis buffer I (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, 2M urea, pH 8.8) and dialysed against the same buffer at 4°C for 12-14 hours. The protein was further dialysed against dialysis buffer II (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Protein concentration was evaluated using the formula:

$$\text{Protein (mg/ml)} = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

### Purification of proteins

To analyse the solubility, pellets obtained from 3.0ml cultures were resuspended in 500 $\mu$ l buffer M1 (PBS pH 7.2). 25 $\mu$ l of lysozyme (10mg/ml) was added and the bacteria incubated for 15 min at 4°C. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and the pellet resuspended in buffer M2 [8M urea, 0.5M NaCl, 20mM imidazole and 0.1M  $NaH_2PO_4$ ] and incubated for 3 to 4 hours at 4°C. After centrifugation, the supernatant was collected and the pellet resuspended in buffer M3 [6M guanidinium-HCl, 0.5M NaCl, 20mM imidazole and 0.1M  $NaH_2PO_4$ ] overnight at 4°C. The

supernatants from all steps were analysed by SDS-PAGE. Some proteins were found to be soluble in PBS, others need urea or guanidium-HCl for solubilization.

For preparative scale purifications, 500ml cultures were induced and fusion proteins solubilized in either buffer M1, M2 or M3 using the procedure described above. Crude extracts were loaded onto a Ni-NTA superflow column (Quiagen) equilibrated with buffer M1, M2 or M3 depending on the solubilization buffer employed. Unbound material was eluted by washing the column with the same buffer. The recombinant fusion protein was eluted with the corresponding buffer containing 500mM imidazole then dialysed against the same buffer in the absence of imidazole.

### **Mice immunisations**

20µg of each purified protein are used to immunise mice intraperitoneally. In the case of some ORFs, Balb-C mice were immunised with Al(OH)<sub>3</sub> as adjuvant on days 1, 21 and 42, and immune response was monitored in samples taken on day 56. For other ORFs, CD1 mice could be immunised using the same protocol. For ORFs 25 and 40, CD1 mice were immunised using Freund's adjuvant, and the same immunisation protocol was used, except that the immune response was measured on day 42, rather than 56. Similarly, for still other ORFs, CD1 mice were immunised with Freund's adjuvant, but the immune response was measured on day 49. Alternatively, 20µg of each purified protein was mixed with Freund's adjuvant and used to immunise CD1 mice intraperitoneally. For many of the proteins, the immunization was performed on days 1, 21 and 35, and immune response was monitored in samples taken on days 34 and 49. For some proteins, the third immunization was performed on day 28, rather than 35, and the immune response was measured on days 20 and 42, rather than 34 and 49.

### **ELISA assay (sera analysis)**

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 7ml of Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10000 rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated

for 2 hours at room temperature and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200 µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200 µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN<sub>3</sub> in PBS) were added to each well and the plates incubated for 90 minutes at 37°C. Wells were washed three times with PBT. 100 µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100 µl of substrate buffer for HRP (25 ml of citrate buffer pH5, 10 mg of O-phenildiamine and 10 µl of H<sub>2</sub>O) were added to each well and the plates were left at room temperature for 20 minutes. 100 µl H<sub>2</sub>SO<sub>4</sub> was added to each well and OD<sub>490</sub> was followed. The ELISA was considered positive when OD<sub>490</sub> was 2.5 times the respective pre-immune sera.

Alternatively, The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10 000rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN<sub>3</sub> in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-



phenildiamine and 10 $\mu$ l of H<sub>2</sub>O<sub>2</sub>) were added to each well and the plates were left at room temperature for 20 minutes. 100 $\mu$ l of 12.5% H<sub>2</sub>SO<sub>4</sub> was added to each well and OD<sub>490</sub> was followed. The ELISA titers were calculated arbitrarily as the dilution of sera which gave an OD<sub>490</sub> value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD<sub>490</sub> of 0.4 was higher than 1:400.

#### **FACScan bacteria Binding Assay procedure.**

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN<sub>3</sub>) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD<sub>620</sub> of 0.07. 100 $\mu$ l bacterial cells were added to each well of a Costar 96 well plate. 100 $\mu$ l of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200 $\mu$ l/well of blocking buffer in each well. 100 $\mu$ l of R-Phicoerytrin conjugated F(ab)<sub>2</sub> goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200 $\mu$ l/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200 $\mu$ l/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

#### **OMV preparations**

Bacteria were grown overnight on 5 GC plates, harvested with a loop and resuspended in 10 ml 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes and the bacteria disrupted by sonication for 10' on ice ( 50% duty cycle, 50% output ). Unbroken cells were removed by centrifugation at 5000g for 10 minutes and the total cell envelope

fraction recovered by centrifugation at 50000g at 4°C for 75 minutes. To extract cytoplasmic membrane proteins from the crude outer membranes, the whole fraction was resuspended in 2% sarkosyl (Sigma) and incubated at room temperature for 20 minutes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, and the supernatant further ultracentrifuged at 50000g for 75 minutes to pellet the outer membranes. The outer membranes were resuspended in 10mM Tris-HCl, pH8 and the protein concentration measured by the Bio-Rad Protein assay, using BSA as a standard.

### **Whole Extracts preparation**

Bacteria were grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30' minutes.

### **Western blotting**

Purified proteins (500ng/lane), outer membrane vesicles (5µg) and total cell extracts (25µg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3 % Tris base, 1.44 % glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

### **Bactericidal assay**

MC58 and 2996 strains were grown overnight at 37°C on chocolate agar plates. 5-7 colonies were collected and used to inoculate 7ml Mueller-Hinton broth. The suspension was incubated at 37°C on a nutator and let to grow until OD<sub>620</sub> was in between 0.5-0.8. The culture was aliquoted into sterile 1.5ml Eppendorf tubes and centrifuged for 20 minutes at maximum speed in a microfuge. The pellet was washed once in Gey's buffer (Gibco) and resuspended in the same buffer to an OD<sub>620</sub> of 0.5, diluted 1:20000 in Gey's buffer and stored at 25°C.

50µl of Gey's buffer/1% BSA was added to each well of a 96-well tissue culture plate. 25µl of diluted (1:100) mice sera (dilution buffer: Gey's buffer/0.2% BSA) were added to each well and the plate incubated at 4°C. 25µl of the previously described bacterial suspension were added to each well. 25µl of either heat-inactivated (56°C waterbath for 30 minutes) or normal baby rabbit complement were added to each well. Immediately after the addition of the baby rabbit complement, 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 0). The 96-well plate was incubated for 1 hour at 37°C with rotation and then 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 1). After overnight incubation the colonies corresponding to time 0 and time 1h were counted.

### Gene Variability

The *ORF4* and *919* genes were amplified by PCR on chromosomal DNA extracted from various *Neisseria* strains (see list of strains). The following oligonucleotides used as PCR primers were designed in the upstream and downstream regions of the genes:

orf 4.1	(forward) CGAATCCGGACGGCAGGACTC
orf 4.3	(reverse) GGCAGGGAATGGCGGATTAAAG
919.1	(forward) AAAATGCCTCTCCACGGCTG or CTGCGCCCTGTGTAAATCCCCT
919.6	(reverse) CAAATAAGAAAGGAATTTTG or GGTATCGCAAACTTCGCCTTAATGCG

The PCR cycling conditions were:

1 cycle	2 min. at 94°
30 cycles	30 sec. at 94°
	30 sec. at ~ 54° or ~ 60° ( in according to T <sub>m</sub> of the primers)
	40 sec. at 72°
1 cycle	7 min. at 72°

The PCR products were purified from 1 % agarose gel and sequenced using the following primers:

orf 4.1 (forward)	CGAATCCGGACGGCAGGACTC
orf 4.2 (forward)	CGACCGCGCCTTTGGGACTG
orf 4.3 (reverse)	GGCAGGGAATGGCGGATTAAAG
orf 4.4 (reverse)	TCTTTGAGTTTGATCCAACC

- 919.1 (forward) AAAATGCCTCTCCACGGCTG or  
CTGCGCCCTGTGTTAAAATCCCCCT
- 919.2 (forward) ATCCTTCCGCCTCGGCTGCG
- 919.3 (forward) AAAACAGCGGCACAATCGAC
- 919.4 (forward) ATAAGGGCTACCTCAAATC
- 919.5 (forward) GCGCGTGGATTATTTTGGG
- 919.6 (reverse) CAAATAAGAAAGGAATTTTG or  
GGTATCGCAAACTTCGCCTTAATGCG
- 919.7 (reverse) CCCAAGGTAATGTAGTGCCG
- 919.8 (reverse) TAAAAAAAGTTCGACAGGG
- 919.9 (reverse) CCGTCCGCCTGTCGTCGCCC
- 919.10 (reverse) TCGTTCCGGCGGGGTCGGGG

All documents cited herein are incorporated by reference in their entireties.

The following Examples are presented to illustrate, not limit, the invention

#### EXAMPLE 1

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

**Table 1: Oligonucleotides used for PCR for Examples 2-10**

ORF	Primer	Sequence	Restriction sites
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT <SEQ ID 3021>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA <SEQ ID 3022>	XhoI
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA <SEQ ID 3023>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC <SEQ ID 3024>	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT <SEQ ID 3025>	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC <SEQ ID 3026>	XhoI
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC <SEQ ID 3027>	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG <SEQ ID 3028>	XhoI
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT <SEQ ID 3029>	BamHI-NdeI

128	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC <SEQ ID 3030>	XhoI
	Forward	CGCGGATCCCATATG-ACTGACAACGCACT <SEQ ID 3031>	BamHI-NdeI
206	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA <SEQ ID 3032>	XhoI
	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA <SEQ ID 3033>	BamHI-NdeI
287	Reverse	CCCGCTCGAG-TTCTGTAAAAAAGTATGTGC <SEQ ID 3034>	XhoI
	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG <SEQ ID 3035>	EcoRI-NheI
406	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC <SEQ ID 3036>	XhoI
	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG <SEQ ID 3037>	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG <SEQ ID 3038>	XhoI

### Localization of the ORFs

The following DNA and amino acid sequences are identified by titles of the following form: [g, m, or a] [#].[seq or pep], where "g" means a sequence from *N. gonorrhoeae*, "m" means a sequence from *N. meningitidis B*, and "a" means a sequence from *N. meningitidis A*; "#" means the number of the sequence; "seq" means a DNA sequence, and "pep" means an amino acid sequence. For example, "g001.seq" refers to an *N. gonorrhoeae* DNA sequence, number 1. The presence of the suffix "-1" to these sequences indicates an additional sequence found for the same ORF, thus, data for an ORF having both an unsuffixed and a suffixed sequence designation applies to both such designated sequences. Further, open reading frames are identified as ORF #, where "#" means the number of the ORF, corresponding to the number of the sequence which encodes the ORF, and the ORF designations may be suffixed with ".ng" or ".a", indicating that the ORF corresponds to a *N. gonorrhoeae* sequence or a *N. meningitidis A* sequence, respectively. The word "partial" before a sequence indicates that the sequence may be a partial or a complete ORF. Computer analysis was performed for the comparisons that follow between "g", "m", and "a" peptide sequences; and therein the "pep" suffix is implied where not expressly stated. Further, in the event of a conflict between the text immediately preceding and describing which sequences are being compared, and the

279 gnm4.seq

1	ATAACGCGGA	TTTGC GGCTG	CTTGATTTC	ACGGTTTTCA	GGGCTTCGGC
51	AAGTTTGTCTG	GCGGCGGGTT	TCATCAGGCT	CGAATGGGAA	GGTACGGACA
101	CGGGCAGCGG	CAGGGCGCGT	TTGGCACCGG	CTTCTTTGGC	GGCAGCCATG
151	GCGCGTCCGA	CGGCGGCGGC	GTTGCCTGCA	ATCACGATTT	GTCCGGGTGA
201	GTTGAAGTTG	ACGGCTTCGA	CCACTTCGCT	TTGGGCGGGCT	TCGGCACAAA
251	TGGCTTTAAC	CTGCTCATCT	TCCAAGCCGA	GAATCGCCGC	CATTGCGCCC
301	ACGCTTGGCG	GTACGGCGGA	CTGCATCAGT	TCGGCGCGCA	GGCGCACGAG
351	TTTGACCGCG	TCGGCAAAAT	TCAATGCGCC	GGCGGCAACG	AGTGCGGTGT
401	ATTGCGCCGAG	GCTGTGTCCG	GCAACGGCGG	CAGGCGTTTT	GCCGCCCCGT
451	TCTAAATAG				

```

1  ITRICGCLIS TVFRASASLS AAGFIRLOWE GTDTGSGRAR LAPASLAAAM
51  ARPTAAALPA ITICPGELKL TASTTSLWAA SAQMALTCSS SKPRIAAIAP
101 TPCGTADCIS SARRRTSLTA SAKFNAPAAT SAVYSPRLCP ATAAGVLP
151 SK*

```

g279.seq

```
1 atgacgcgga tttgcggctg cttgatttca acggttttga gtgtttcggc
51 aagtttgctg gcggcggtt tcatcaggct gcaatgggaa ggaacggata
101 ccggcagcgg cagggcgctg ttggctccgg cttctttggc ggcagccatg
151 gtgcgtccga cggcgcgggc gttgcctgca atcacgactt gtccgggcga
201 gttgaagttg acggcttcga ccacttcgcc ctgtgcggat tcggcacaaa
251 tctgcctgac ctgttcatct tccaaaccca aaatggccgc cattgcgcct
301 acgccttgcg gtacggcgga ctgcatcagt tcggcgcgca ggcggacgag
351 tttgacggga tcggcaaaat ccaatgcttc ggcggcgaca agcgcggtgt
401 attcgccgag gctgtgtcgg gcaacggcgg caggcgtttt gccgccact
451 tccaaatag
```

g279.pep

1	<u>MTRICGCLIS</u>	<u>TVLSVSASLS</u>	<u>AAGFIRLOWE</u>	<u>GTDTGSGRAR</u>	<u>LAPASLAAAM</u>
51	<u>VRPTAAALPA</u>	<u>ITTCPGELKL</u>	<u>TASTTSPCAD</u>	<u>SAQICLTCSS</u>	<u>SKPKMAAIAP</u>
101	<u>TPCGTADCIS</u>	<u>SARRRTSLTA</u>	<u>SAKSNASAAT</u>	<u>SAVYSPRLCP</u>	<u>ATAAGVLPPT</u>
151	SK*				

```

      10          20          30          40          50          60
m279.pep ITRICGCLISTVFRASASLSAAGFIRLQWEGTDTGSGRRARLAPASLAAAMARPTAAALPA
           :| |: | |: | |: | |: | |: | |: | |: | |: | |: | |: | |: |
g279     MTRICGCLISTVLSVASLSAAGFIRLQWEGTDTGSGRRARLAPASLAAAMVRPTAAALPA
           10          20          30          40          50          60

```



```

              70          80          90          100         110         120
m279.pep      ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA
               ||||||||| | : ||||| : : |||||||
g279           ITTCPGELKLTASTTSPCADSAQICLTCSSSKPKMAAIAPTPCGTADCISSARRRTSLTA
               70          80          90          100         110         120

              130        140        150
m279.pep      SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX
               ||| ||||||| : |||
g279           SAKSNSAATSAVYSPRLCPATAAGVLPPTSKX
               130        140        150
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3043>:

a279.seq

1	ATGACNCNGA	TTTGC	GGCTG	CTTGATT	CA	ACGGTTT	NNA	GGGCTT	CGGC
51	GAGTTTGT	CG	GCGGCGGGT	TCATGAGG	CT	GCAATGGG	AA	GGTACN	GACA
101	CNNGCAG	CGG	CAGGGCGCGT	TTGGCGCC	GG	CTTCTTT	GGC	GGCAAG	CATA
151	GCGCGCT	CGA	CGGCGCGCGG	ATTGCTT	GCA	ATCACGAC	TT	GTCGCG	GCGA
201	GTTGAAG	TTG	ACGGCTT	CA	CCACTT	CATC	CTGTGCG	GAT	TCCGCG
251	TTTGT	TTTAC	CTGTT	CATCT	TCCAAG	CCGA	GAATCG	CCGC	CATTGC
301	ACGCCTT	GCG	GTACGG	CGGA	CTGCAT	CAGT	TGGCG	CGCA	NGCGAC
351	TTTGAC	CCGCG	TCCGCA	AAAAT	CCAATG	CGCC	GGCGG	CAACN	AGTGC
401	ATTGCC	CGAN	GCTGT	GTCG	G	GCAACG	CGCG	CAGGCG	TTTT
451	TCCGAAT	AG						GCCGCC	CGCT

This corresponds to the amino acid sequence <SEQ ID 3044; ORF 279.a>:

a279.pep

1	<u>MTXICGCLIS</u>	<u>TVXRASASLS</u>	<u>AAGFMRLQWE</u>	<u>GTDTGSGRAR</u>	<u>LAPASLAASI</u>
51	ARSTAAALPA	ITTCGELKL	TASTTSSCAD	SAQICFTCSS	SKPRIAAIAP
101	TPCGTADCIS	SARXRSLTA	SAKSNAPEAT	SAVYSPXLCP	ATAAGVLPPA
151	SE*				

**m279/a279** ORFs 279 and 279.a showed a 88.2% identity in 152 aa overlap

	10	20	30	40	50	60
m279.pep	ITRICGCLISTVFRASASLSAAGFIRLOWEGTDTGSGRARLAPASLAAAMARPTAAALPA :          :          :          :          :					
a279	MTXICGCLISTVXRASASLSAAGFMRLQEWEGTDTGSGRARLAPASLAASIARSTAAALPA 10 20 30 40 50 60					
	70	80	90	100	110	120
m279.pep	ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA                   :      :					
a279	ITTCPGELKLTASTTSSCADSAQICFTCSSSKPRIAAIAPTPCGTADCISSARXRTSLTA 70 80 90 100 110 120					
	130	140	150			
m279.pep	SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX 					
a279	SAKSNAPAATSAVYSPXLCPATAAGVLPPASEX 130 140 150					

519 and 519-1 gnm7.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3045>:

m519.seq (partial)

1	..TCCGTTATCG	GGCGTATGGA	GTTGGACAAA	ACGTTTGAAG	AACGCGACGA
51	AATCAACAGT	ACTGTTGTTG	CGGCTTTGGA	CGAGGCGGCC	GGGgCTTgGG
101	GTGTGAAGST	TTTGCGTTAT	GAGATTAAAG	ACTTGGTTCC	GCCCGCAAGAA
151	ATCCTTCGCT	CAATCGCAGC	GCAAATTACT	GCCGAACGCG	AAAAAACGCG
201	CCGTATCGCC	GAATCCGAAG	GTCGTAATAA	TGCAACAATC	AACCTTGCCA

```

251 GTGGTCAGCG CGAAGCCGAA ATCCAACAAT CCGAAGGCGA GGCTCAGGCT
301 GCGGTCAATG CGTCAAATGC CGAGAAAATC GCCCGCATCA ACCGCGCCAA
351 AGGTGAAGCG GAATCCTTGC GCCTTGTTGC CGAAGCCAAT GCCGAAGCCA
401 TCCGTCAAAT TGCCGCGGCC CTTCAAACCC AAGGCGGTGC GGATGCGGTC
451 AATCTGAAGA TTGCGGAACA ATACGTCGCT GCGTTCAACA ATCTTGCCAA
501 AGAAAGCAAT ACGCTGATTA TGCCCGCCAA TGTGCGGAC ATCGGCAGCC
551 TGATTTCTGC CGGTATGAAA ATTATCGACA GCAGCAAAAC CGCCAAaTAA

```

This corresponds to the amino acid sequence <SEQ ID 3046; ORF 519>:

```

m519.pep (partial)
1 ..SVIGRMELDK TFEERDEINS TVVAALDEAA GAWGVKVLRY EIKDLVPPQE
51 ILRSMQAQIT AEREKRARIA ESEGRKIEQI NLASGQREAE IQQSEGEAQA
101 AVNASNAEKI ARINRAKGEA ESLRLVAEAN AEATRQIAAA LQTGGGADAV
151 NLKIAEQYVA AFNNLAKESN TLIMPANVAD IGSLISAGMK IIDSSKTAK*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3047>:

```

g519.seq
1 atggaatttt tcattatctt gttggcagcc gtcgccgttt tcggcttcaa
51 atcctttgtc gtcatecccc agcaggaagt ccacgttggtc gaaaggctcg
101 ggcgtttcca tcgcgccttg acggcgggtt tgaatatattt gattcccttt
151 atcgaccgcg tcgcctaccg ccattcgctg aaagaaatcc ctttagacgt
201 acccagccag gtctgcatca cgcgcgataa tacgcaattg actgttgacg
251 gcatcatcta tttccaagta accgatccca aactgccttc atacggttcg
301 agcaactaca ttatggcaat taccagctt gcccaaacga cgctgcgttc
351 cgttatcggt cgatggagt tggacaaaac gtttgaagaa cgcgacgaaa
401 tcaacagtac cgtcgtctcc gccctcgatg aagccgcccg ggcttggggg
451 gtgaaagtcc tccgttacga aatcaaggat ttggttccgc cgcaagaaat
501 ccttcgcgca atgcaggcac aaattaccgc cgaacgcgaa aaacgcgccc
551 gtattgccga atccgaaggc cgtaaaatcg aacaaatcaa ctttgccagt
601 ggtcagcgtg aagccgaaat ccaacaatcc gaaggcgagg ctcaggctgc
651 ggtcaatgcg tccaatgccg agaaaatcgc ccgcatcaac cgcgccaaag
701 gcgaagcgga atccctgcgc cttgttgccg aagccaatgc cgaagccaac
751 cgtcaaattg ccgccgccct tcaaacccaa agcggggcgg atgcggtcaa
801 tctgaagatt gcgggacaat acgttaccgc gttcaaaaat cttgccaaag
851 aagacaatac gcggattaag cccgccaaag ttgcccgaat cgggaacctt
901 aattttcggc ggcatgaaaa attttcgcca gaagcaaaaa cggccaaata
951 a

```

This corresponds to the amino acid sequence <SEQ ID 3048; ORF 519.ng>:

```

g519.pep
1 MEFFIILLAA VAVFGFKSFV VIPQEVHV V ERLGRFHRAL TAGLNILIPF
51 IDRVAYRHSI KEIPLDVPSQ VCITRDNTQL TVDGIYFQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEAQAANA SNAEKIARIN RAKGEAESLR LVAEANAEAN
251 RQIAAALQTQ SGADAVNLKI AGQYVTAFKN LAKEDNTRIK PAKVAEIGNP
301 NFRRHEKFSF EAKTAK*

```

ORF 519 shows 87.5% identity over a 200 aa overlap with a predicted ORF (ORF 519.ng) from *N. gonorrhoeae*:

```

m519/g519

m519.pep                               10      20      30
SVIGRMELDKTFEERDEINSTVVAALDEAA
g519      YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
          90      100     110     120     130     140

m519.pep          40      50      60      70      80      90
GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE

```



```

g519      |||||:|||||
          GAWGVKVLRYEIKDLVPPQEILRAMQAQITAEREKRARIAESEGKIEQINLASGQREAE
          150      160      170      180      190      200

          100      110      120      130      140      150
m519.pep  IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAEAIRQIAAALQTQGGADAV
          |||||:|||||
g519      IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAEANRQIAAALQTQSGADAV
          210      220      230      240      250      260

          160      170      180      190      200
m519.pep  NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL-ISAGMKIIDSSKTAK
          |||||:|||||
g519      NLKIAGQYVTAFFKNLAKEDNTRIKPAKVAEIGNPNFRRHEKFSPEAKTAK
          270      280      290      300      310

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3049>:

```

a519.seq
1  ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA
51  ATCCTTTGTT GTCATCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GCGTTTTCCA TCGCGCCCTG ACGGCCGGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GTATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTTCG
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGCAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG AGCTTGGGGT
451 GTGAAGGTTT TCGGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC TGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAAATGCCG AGAAAATCG CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCCGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAT TATCGACAGC AGCAAAACCG CCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 3050; ORF 519.a>:

```

a519.pep
1  MEFFFIILLAA VVVFGEKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
51  IDRVAYRHSI KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLSYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 QREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

**m519/a519** ORFs 519 and 519.a showed a 99.5% identity in 199 aa overlap

```

m519.pep
          10      20      30
          SVIGRMELDKTFEERDEINSTVVAALDEAA
          |||||:|||||
a519      YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
          90      100      110      120      130      140

          40      50      60      70      80      90
m519.pep  GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGKIEQINLASGQREAE
          |||||:|||||
a519      GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGKIEQINLASGQREAE
          150      160      170      180      190      200

          100      110      120      130      140      150

```

```

m519.pep      IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAIAIRQIAAALQTQGGADAV
a519          |||||
               210      220      230      240      250      260
m519.pep      NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSSKTAKX
a519          |||||
               270      280      290      300      310

```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3051>:

```

m519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGTAGCC GTCGCCGTTT TCGGTTTCAA
51 ATCCTTTGTT GTCATCCAC AACAGGAAGT CCACGTTGTC GAAAGGCTGG
101 GCGGTTTCCA TCGCGCCCTG ACGGCGGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTGCTG AAAGAAATCC CTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTTCG
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC TGTTGTTGCG GCTTTGGACG AGGCGGCCGG GGCTTGGGGT
451 GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCGCG CGCAAGAAAT
501 CCTTCGCTCA ATGCAAGCGC AAATTACTGC CGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCTGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAT TATCGACAGC AGCAAAACCG CCAAAATA

```

This corresponds to the amino acid sequence <SEQ ID 3052; ORF 519-1>:

```

m519-1.
1  MEFFIILLVA VAVFGFKSFV VIPQQEVHV ERLGRFHRAL TAGLNILIPF
51 IDRVAIRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVA ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEEAQAVNA SNAEKIARIN RAKGEAESLR LVAEANAIAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3053>:

```

g519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGCAGCC GTCGCCGTTT TCGGCTTCAA
51 ATCCTTTGTC GTCATCCCCC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GCGGTTTCCA TCGCGCCCTG ACGGCCGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTGCTG AAAGAAATCC CTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGATAA TACGCAATTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGATCCCA AACTCGCCTC ATACGGTTTCG
301 AGCAACTACA TTATGGCAAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC CGTCGCTCTC GCCCTCGATG AAGCCGCCGG GGCTTGGGGT
451 GTGAAGTCC TCCGTTACGA AATCAAGGAT TTGGTTCGCG CGCAAGAAAT
501 CCTTCGCGCA ATGCAAGCAC AAATTACCGC CGAACGCGAA AAACGCGCCC
551 GTATTGCCGA ATCCGAAGGC CGTAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGTG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCATGCG TCCAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GCGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCGCCCT TCAAACCCAA GGCGGGGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTAGCCGC GTTCAACAAT CTTGCCAAAG

```

851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG  
901 ATTTCTGCCG GCATGAAAT TATCGACAGC AGCAAAACCG CCAAATAA

This corresponds to the amino acid sequence <SEQ ID 3054; ORF 519-1.ng>:

g519-1.pep  
1 MEFFFIILLAA VAVFGFKSFV VIPQQEVHV V ERLGRFHRAL TAGLNILIPF  
51 IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKCLASYGS  
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG  
151 VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS  
201 GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAETI  
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL  
301 ISAGMKIIDS SKTAK\*

m519-1/g519-1 ORFs 519-1 and 519-1.ng showed a 99.0% identity in 315 aa overlap

g519-1.pep	10	20	30	40	50	60
m519-1	10	20	30	40	50	60
g519-1.pep	70	80	90	100	110	120
m519-1	70	80	90	100	110	120
g519-1.pep	130	140	150	160	170	180
m519-1	130	140	150	160	170	180
g519-1.pep	190	200	210	220	230	240
m519-1	190	200	210	220	230	240
g519-1.pep	250	260	270	280	290	300
m519-1	250	260	270	280	290	300
g519-1.pep	310					
m519-1	310					

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3055>:

a519-1.seq  
1 ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA  
51 ATCCTTTGTT GTCATCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG  
101 GCGGTTTCCA TCGCGCCCTG ACGGCCGTTT TGAATATTTT GATTCCCTTT  
151 ATCGACCGCG TCGCCTACCG CCATTTCGTC AAAGAAATCC CTTTAGACGT  
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCGCTG ACTGTTGACG  
251 GTATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTTCG  
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC  
351 CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA

401	TCAACAGCAC	CGTCGCTCC	GCCCTCGATG	AAGCCGCCGG	AGCTTGGGGT
451	GTGAAGGTTT	TGCGTTATGA	GATTAAAGAC	TTGGTTCGCG	CGCAAGAAAT
501	CCTTCGCTCA	ATGCAGGCGC	AAATTACTGC	TGAACGCGAA	AAACGCGCCC
551	GTATCGCCGA	ATCCGAAGGT	CGTAAATCG	AACAAATCAA	CCTTGCCAGT
601	GGTCAGCGCG	AAGCCGAAAT	CCAAACAATCC	GAAGGCGAGG	CTCAGGCTGC
651	GGTCAATGCG	TCAAATGCCG	AGAAAAATCC	CCGCATCAAC	CGCGCCAAAG
701	GTGAAGCGGA	ATCCTTGCGC	CTTGTTCGCG	AAGCGATGCG	CGAAGCCATC
751	CGTCAAATTG	CCGCGGCCCT	TCAAAACCAA	GGCGGTGCGG	ATGCGGTCOA
801	TCTGAAGATT	GCGGAACAAT	ACGTCGCCGC	GTTCAACAAT	CTTGCCAAAG
851	AAAGCAATAC	GCTGATTATG	CCCGCCAATG	TTGCCGACAT	CGGCAGCCTG
901	ATTTCTGCCG	GTATGAAAAT	TATCGACAGC	AGCAAAACCG	CCAAATAA

This corresponds to the amino acid sequence <SEQ ID 3056; ORF 519-1.a>:

a519-1.pcp.

1	MEFFIILLAA	VVVFGEKSFV	VIPOQEVHV	ERLGRFHRAL	TAGNLILIPF
51	IDRVAYRHS	KEIPLDVPSQ	VCITRDNTQL	TVDGIIYFQV	TDPKLSYSGS
101	SNYIMAITQL	AQTTLSVIG	RMELDKTFEE	RDEINSTVVS	ALDEAAGAWG
151	VKVLRYEIKD	LVPQEILRS	MQAQITAERE	KRARIAESEG	RKIEQINLAS
201	GOREAAIQQS	EGEAQAAVNA	SNAEKIARIN	RAKGEAESLR	LVAEANAIAI
251	QRIAAALQTQ	GKADAVNLKI	AEQYVAAFNN	LAKESNTLIM	PANVADIGSL
301	ISAGMKIIDS	SKTAP*			

**m519-1/a519-1** ORFs 519-1 and 519-1.a showed a 99.0% identity in 315 aa overlap

a519-1.pep	MEFFFIILLAAVVVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL	10	20	30	40	50	60
m519-1	MEFFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL	10	20	30	40	50	60
a519-1.pep	KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG	70	80	90	100	110	120
m519-1	KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG	70	80	90	100	110	120
a519-1.pep	RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEIILRSMQAQITAERE	130	140	150	160	170	180
m519-1	RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEIILRSMQAQITAERE	130	140	150	160	170	180
a519-1.pep	KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR	190	200	210	220	230	240
m519-1	KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR	190	200	210	220	230	240
a519-1.pep	LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL	250	260	270	280	290	300
m519-1	LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL	250	260	270	280	290	300
a519-1.pep	ISAGMKIIDSSKTAKX	310					
m519-1	ISAGMKIIDSSKTAKX	310					

m576.seq.. (partial)

1	..ATGCAGCAGG	CAAGCTATGC	GATGGGCGTG	GACATCGGAC	GCTCCCTGAA
51	GCAAATGAAG	GAACAGGGCG	CGGAAATCGA	TTTGAAAGTC	TTTACCGAAG
101	CCATCGCAGGC	AGTGTATGAC	GGCAAAGAAA	TCAAATGAC	CGAAGAGCAG
151	GCTCAGGAAG	TCATGATGAA	ATTCTTTCAG	GAACAACAGG	CTAAAGCCGT
201	AGAAAAACAC	AAGCGGGACG	ACAGGCCCAA	TAAAGAAAA	GGCGAAGCCCT
251	TTCTGAAAGA	AAATGCCGCC	AAAGACGGCG	TGAAGACCAC	TGCTTCCGGC
301	CTGCAATACA	AAATCACCAA	ACAGGGCGAA	GGCAAACAGC	CGACCAAAGA
351	CGCATCGTT	ACCGTGGAAT	ACGAAGGCCG	CCTGATTGAC	GGTACGGTAT
401	TCGACAGCAG	CAAAGCCAA	GCGGCCCCGG	TCACCTTCCC	TTTGAGCCAA
451	GTGATTCCGG	GTTGGACCGA	AGCGCTACAG	CTTCTGAAAG	AAGGCGCCGA
501	AGCCACGTT	TACATCCCGT	CCAACCTTGC	CTACCGCGAA	CAGGGTGCGG
551	GCGACAAAAT	CGGTCCGAAC	GCCACTTTGG	TATTTGATGT	GAAACTGGTC
601	AAAATCGGCG	CACCCGAAAA	CGCGCCCCGC	AAGCAGCCCG	CTCAAGTCGA
651	CATCAAAAAA	GTAATAAT			

m576.pap.. (partial)

```

1      .MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
51     AQEVMMKFLQ  EQQAKAVEKH KADAKANKEK GEAFLENAA  KDGVKTTASG
101    LQYKITKQGE  GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
151    VIPGWTEGV  LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVDVKLV
201    KIGAPENAPA  KOPAOVNDIKK VN*

```

q576.seq.. (partial)

1	..atgggcgtgg	acatcggacg	ctccctgaaa	caaatgaagg	aacaggggcgc
51	ggaaatcgat	ttgaaagtct	ttaccgatgc	catgcaggca	gtgtatgacg
101	gcaaaagaat	caaaatcgacc	gaagagcagg	cccaggaagt	gatgatgaaa
151	ttcctgcagg	agcagcaggc	taaaggcgta	gaaaaaacaca	aggcggatgc
201	gaaggccaac	aaagaaaaag	cgaaagcctt	cgttaaggaa	aatgccgcgc
251	aagacggcgt	gaagaccact	gcttccggtc	tgcagtacaa	aatcaccaaa
301	cagggtgaag	gcaaacagcc	gacaaaagac	gacatcgcta	ccgtggaata
351	cgaaggccgc	ctgattgacg	gtaccgtatt	cgacagcagc	aaagccaacg
401	gcggcccggc	caccttcctt	ttgaaggcaag	tgattccggg	ttggaccgaa
451	ggcgtacggc	ttctgaaaga	aggcggcgaa	gccacgttct	acatcccgtc
501	caaccttgcc	taccgcgaac	agggtgcggg	cgaaaaaatc	ggtcggaacg
551	ccactttggt	atttgacgtg	aaactgggtca	aaatcggcgc	accggaaaac
601	gcgcccgcga	agcagccgga	tcaagtcqac	atcaaaaaag	taaattaa

g576.pep.. (partial)

```

1  ..MGVDIGRSLK QMKEQGAIED LKVFTDAMQA VYDGKEIKMT EEQAQEVMMK
51  FLQEQQAKAV EKHKADAKAN KEKGEAFLKE NAAEDGVKTT ASGLQYKITK
101 QGEGKQPTKD DIVTVEYEGR LIDGTVFDSS KANGGPATFP LSQVIPGWTE
151 GVRLLKEGGE ATFYIPSNLA YREQGAGEKI GPNATLVFDV KLVKIGAPEN
201 APAKOPDOVD IKKVN*

```

### Homology with a predicted ORF from *N. gonorrhoeae*

**m576/g576** ORFs 576 and 576.ng showed a 97.2% identity in 215 aa overlap

m576.pep                    10                    20                    30                    40                    50                    60  
 MQQASYAMGVDIGRSLKMKKEQGAIEDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQ  
 |||||:|||||

```

g576          MGVDIGRSLKQMKEQGAEIDLKVFTDAMQAVYDGKEIKMTEEQAEVMMKFLQ
                10      20      30      40      50
m576.pep      70      80      90      100     110     120
EQQAKAVEKHKADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQEGEKQPTKDDIV
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          60      70      80      90      100     110
EQQAKAVEKHKADAKANKEKGEAFLKENAAEDGVKTTASGLQYKITKQEGEKQPTKDDIV

m576.pep      130     140     150     160     170     180
TVEYEGRLIDGTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYRE
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          120     130     140     150     160     170
TVEYEGRLIDGTVFDSSKANGGPATFPLSQVIPGWTEGVRLLEKGEATFYIPSNLAYRE

m576.pep      190     200     210     220
QGAGDKIGPNATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          180     190     200     210
QGAGEKIGPNATLVFDVKLVKIGAPENAPAKQPDQVDIKKVN

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3061>:

```

a576.seq
1   ATGAACACCA TTTTCAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
51  ACTTTCCGCC TCGCGCAAAA AAGAAGCCGC CCCGCATCT GCATCCGAAC
101 CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCAGC
151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201 GCAAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG
301 GCTCAGGAAG TCATGATGAA ATTCTTCAG GAACAACAGG CTAAAGCCGT
351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGCC
451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
601 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
701 GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3062; ORF 576.a>:

```

a576.pep
1   MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASSA QGDTSSIGST
51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
101 AQEVMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*

```

m576/a576 ORFs 576 and 576.a showed a 99.5% identity in 222 aa overlap

```

m576.pep      10      20      30
                MQQASYAMGV DIGRSLKQMK EQGAEIDLKV
a576          30      40      50      60      70      80
CGKKEAAPASASEPAASSAQGDTSSIGSTMQQASYAMGV DIGRSLKQMK EQGAEIDLKV

m576.pep      40      50      60      70      80      90
FTEAMQAVYDGKEIKMTEEQAEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a576          90      100     110     120     130     140
FTEAMQAVYDGKEIKMTEEQAEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA

```

	100	110	120	130	140	150
m576.pep	KDGVKTTASGLQYKITKQEGKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ					
a576	KDGVKTTASGLQYKITKQEGKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ					
	150	160	170	180	190	200
	160	170	180	190	200	210
m576.pep	VIPGWTEGVQLLEGGGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA					
a576	VILGWTEGVQLLEGGGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA					
	210	220	230	240	250	260
	220					
m576.pep	KQPAQVDIKKVN					
a576	KQPAQVDIKKVN					
	270					

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3063>:

m576-1.seq

1	ATGAACACCA	TTTTCAAAT	CAGCGCACTG	ACCCTTTCCG	CCGCTTTGGC
51	ACTTTCCGCC	TGCGGCAAAA	AAGAAGCCGC	CCCCGCATCT	GCATCCGAAC
101	CTGCCGCCGC	TTCTTCCGCG	CAGGGCGACA	CCTCTTCGAT	CGGCAGCAGC
151	ATGCAGCAGG	CAAGCTATGC	GATGGGCGTG	GACATCGGAC	GCTCCCTGAA
201	GCAAATGAAG	GAACAGGGCG	CGGAAATCGA	TTTGAAAGTC	TTTACCGAAG
251	CCATGCAGGC	AGTGTATGAC	GGCAAAGAAA	TCAAATGAC	CGAAGAGCAG
301	GCTCAGGAAG	TCATGATGAA	ATTCCTTCAG	GAACAACAGG	CTAAAGCCGT
351	AGAAAAACAC	AAGGCGGACG	CGAAGGCCAA	TAAAGAAAAA	GGCGAAGCCT
401	TTCTGAAAGA	AAATGCCGCC	AAAGACGGCG	TGAAGACCAC	TGCTTCCGGC
451	CTGCAATACA	AAATCACCAA	ACAGGGCGAA	GGCAAACAGC	CGACCAAAGA
501	CGACATCGTT	ACCGTGGAAT	ACGAAGGCCG	CCTGATTGAC	GGTACGGTAT
551	TCGACAGCAG	CAAAGCCAAC	GGCGGCCCGG	TCACCTTCCC	TTTGAGCCAA
601	GTGATTCCGG	GTTGGACCGA	AGGCGTACAG	CTTCTGAAAG	AAGGCGGCGA
651	AGCCACGTTT	TACATCCCGT	CCAACCTTGC	CTACCGCGAA	CAGGGTGCGG
701	GCGACAAAAT	CGGTCCGAAC	GCCACTTTGG	TATTTGATGT	GAAACTGGTC
751	AAAATCGGCG	CACCCGAAAA	CGCGCCCGCC	AAGCAGCCGG	CTCAAGTCGA
801	CATCAAAAAA	GTAAATTAA			

This corresponds to the amino acid sequence <SEQ ID 3064; ORF 576-1>:

m576-1.pep

1	MNTIFKISAL	TLAALALSA	CGKKEAAPAS	ASEPAAASSA	QGDTSSIGST
51	MQQASYAMGV	DIGRSLKQMK	EQGAEIDLKV	FTEAMQAVYD	GKEIKMTEEQ
101	AQEVMMKFLO	EQQAKAVEKH	KADAKANKEK	GEAFLKENAA	KDGVKTTASG
151	LQYKITKQGE	GKQPTKDDIV	TVEYEGRLID	GTVPDSSKAN	GGPVTFPLSQ
201	VIPGWTEGVQ	LLKEGGGEATF	YIPSNLAYRE	QGAGDKIGPN	ATLVFDVKLV
251	KIGAPENAPA	KQPAQVDIKK	VN*		

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3065>:

g576-1.seq

1	ATGAACACCA	TTTTCAAAT	CAGCGCACTG	ACCCTTTCCG	CCGCTTTGGC
51	ACTTTCCGCC	TGCGGCAAAA	AAGAAGCCGC	CCCCGCATCT	GCATCCGAAC
101	CTGCCGCCGC	TTCTGCCGCG	CAGGGCGACA	CCTCTTCAAT	CGGCAGCAGC
151	ATGCAGCAGG	CAAGCTATGC	AATGGGCGTG	GACATCGGAC	GCTCCCTGAA
201	ACAAATGAAG	GAACAGGGCG	CGGAAATCGA	TTTGAAAGTC	TTTACCGATG
251	CCATGCAGGC	AGTGTATGAC	GGCAAAGAAA	TCAAATGAC	CGAAGAGCAG
301	GCCCAGGAAG	TGATGATGAA	ATTCCTGCAG	GAGCAGCAGG	CTAAAGCCGT
351	AGAAAAACAC	AAGGCGGATG	CGAAGGCCAA	CAAAGAAAAA	GGCGAAGCCT
401	TCCTGAAGGA	AAATGCCGCC	AAAGACGGCG	TGAAGACCAC	TGCTTCCGGT
451	CTGCAGTACA	AAATCACCAA	ACAGGGTGAA	GGCAAACAGC	CGACAAAAGA
501	CGACATCGTT	ACCGTGGAAT	ACGAAGGCCG	CCTGATTGAC	GGTACCGTAT
551	TCGACAGCAG	CAAAGCCAAC	GGCGGCCCGG	CCACCTTCCC	TTTGAGCCAA
601	GTGATTCCGG	GTTGGACCGA	AGGCGTACGG	CTTCTGAAAG	AAGGCGGCGA
651	AGCCACGTTT	TACATCCCGT	CCAACCTTGC	CTACCGCGAA	CAGGGTGCGG

701 GCGAAAAAAT CGGTCCGAAC GCCACTTTGG TATTTGACGT GAAACTGGTC  
751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG ATCAAGTCGA  
801 CATCAAAAAA GTAAATTAA

This corresponds to the amino acid sequence <SEQ ID 3066; ORF 576-1.ng>:

g576-1.pep

1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASAA QGDTSSIGST  
51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTDAMQAVYD GKEIKMTEEQ  
101 AQEVMMKFLO EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVTKTASG  
151 LQYKITQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPATFPLSQ  
201 VIPGWTEGVR LLKEGGEATF YIPSNLAYRE QGAGEKIGPN ATLVFDVKLV  
251 KIGAPENAPA KQPDQVDIKK VN\*

g576-1/m576-1  
overlap

ORFs 576-1 and 576-1.ng showed a 97.8% identity in 272 aa

g576-1.pep	10	20	30	40	50	60
	MNTIFKISAL	TLSAALALSA	SACGKKEAAPAS	ASEPAASAA	QGDTSSIGST	MQQASYAMGV
m576-1	10	20	30	40	50	60
	MNTIFKISAL	TLSAALALSA	SACGKKEAAPAS	ASEPAASAA	QGDTSSIGST	MQQASYAMGV
g576-1.pep	70	80	90	100	110	120
	DIGRSLKQMK	EQGAEIDLKV	FTDAMQAVYD	GKEIKMTEE	QAQEVMMKF	LQEQQAKAVEKH
m576-1	70	80	90	100	110	120
	DIGRSLKQMK	EQGAEIDLKV	FTDAMQAVYD	GKEIKMTEE	QAQEVMMKF	LQEQQAKAVEKH
g576-1.pep	130	140	150	160	170	180
	KADAKANKEK	GEAFLENAA	KDGVTKTAS	GLQYKITQ	GEGKQPTK	DDIVTVEYEGRLID
m576-1	130	140	150	160	170	180
	KADAKANKEK	GEAFLENAA	KDGVTKTAS	GLQYKITQ	GEGKQPTK	DDIVTVEYEGRLID
g576-1.pep	190	200	210	220	230	240
	GTVFDSSKAN	GGPATFPLS	QVIPGWTEG	VRLLEKEG	GEATFYIP	SNLAYREQGAGEKIGPN
m576-1	190	200	210	220	230	240
	GTVFDSSKAN	GGPATFPLS	QVIPGWTEG	VRLLEKEG	GEATFYIP	SNLAYREQGAGEKIGPN
g576-1.pep	250	260	270			
	ATLVFDVKLV	KIGAPENAP	AKQPDQVD	IKKVN		
m576-1	250	260	270			
	ATLVFDVKLV	KIGAPENAP	AKQPDQVD	IKKVN		

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3067>:

a576-1.seq

1 ATGAACACCA TTTTCAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC  
51 ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCC GCATCCGAAC  
101 CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCAGC  
151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCCGAC GCTCCCTGAA  
201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG  
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG  
301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT  
351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT  
401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC  
451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAGA  
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT  
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA  
601 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTCTGAAAG AAGGCGGCGA  
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG



701 GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC  
 751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA  
 801 CATCAAAAAA GTAAATTAA

This corresponds to the amino acid sequence <SEQ ID 3068; ORF 576-1.a>:

a576-1.pep

1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASSA QGDTSSIGST  
 51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ  
 101 AQEVMMKFLO EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG  
 151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ  
 201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV  
 251 KIGAPENAPA KQPAQVDIKK VN\*

a576-1/m576-1

ORFs 576-1 and 576-1.a showed a 99.6% identity in 272 aa overlap

a576-1.pep	10	20	30	40	50	60
	MNTIFKISALTLSAALALSACGKKEAAPASASEPAASSAQGDTSSIGSTMQQASYAMGV					
m576-1	10	20	30	40	50	60
	MNTIFKISALTLSAALALSACGKKEAAPASASEPAASSAQGDTSSIGSTMQQASYAMGV					
a576-1.pep	70	80	90	100	110	120
	DIGRSLKQMKEQGAEIDLKVFTEAMQAVYDQKEIKMTEEQAEVMMKFLOEQQAKAVEKH					
m576-1	70	80	90	100	110	120
	DIGRSLKQMKEQGAEIDLKVFTEAMQAVYDQKEIKMTEEQAEVMMKFLOEQQAKAVEKH					
a576-1.pep	130	140	150	160	170	180
	KADAKANKEKGEAFLENAAKDGVKTTASGLQYKITKQEGGKQPTKDDIVTVEYEGRLID					
m576-1	130	140	150	160	170	180
	KADAKANKEKGEAFLENAAKDGVKTTASGLQYKITKQEGGKQPTKDDIVTVEYEGRLID					
a576-1.pep	190	200	210	220	230	240
	GTVFDSSKANGGPVTFPLSQVILGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN					
m576-1	190	200	210	220	230	240
	GTVFDSSKANGGPVTFPLSQVILGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN					
a576-1.pep	250	260	270			
	ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN					
m576-1	250	260	270			
	ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN					

919 gnm43.seq

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3069>:

m919.seq

1 ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TACGGCATCG CCGCCGCCAT  
 51 CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA  
 101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC  
 151 GGAACGACGG TCGGCGGCGG CGGGGCCGTC TATACCGTTG TACCGCACCT  
 201 GTCCCTGCCC CACTGGGCGG CGCAGGATT CGCCAAAAGC CTGCAATCCT  
 251 TCCGCCTCGG CTGCGCCAAT TTGAAAACC GCCAAGGCTG GCAGGATGTG  
 301 TGCGCCAAG CCTTTCAAAC CCCCCTCCAT TCCTTTCAGG CAAAACAGTT  
 351 TTTTGAACGC TATTTACGCG CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG  
 401 CCGGTACGGT TACCGGCTAT TACGAACCGG TGCTGAAGGG CGACGACAGG

```

451 CGGACGGCAC AAGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCCTGCCG GTTTCGGGAG CGGAAAAGCC CTTGTCCGCA
551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
601 CATACGCGCG ACCTCTCCCG ATTCCCATC ACCGCGCGCA CAACAGCAAT
651 CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC
701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCATACT CGGTTACGCC
751 GAAGACCTTG TCGAATTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
851 AACATCCyTA CGTTTCCATC GGACGTATA TGGCGGATAA GGGCTACCTC
901 AAATCTGGAC AAACCTCCAT GCAGGGCATT AAGTCTTATA TGCGGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC AGCAATGACG GCCCTGTCGG CGCACTGGGC
1051 ACGCGCTGA TGGGGGAATA TGCCGCGCA GTCGACCGGC ACTACATTAC
1101 CTTGGGTGCG CCCTTATTTG TCGCCACCGC CCATCCGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTC CTACCCAACG
1301 GTATGAAGCC CGAATACCGC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3070; ORF 919>:

m919.pep

```

1  MKKYLFRAL YGIAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
51  GTTVGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SFQAKOFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR
151 RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT
201 HTADLSRFPI TARTTAIKGR FEGRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IOGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KSYMQRNPQR LAEVLGQNPS YIFFRELAGS SNDGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 3071>:

g919.seq

```

1  ATGAAAAAAC ACCTGCTCCG CTCCGCCCTG TACGGcatCG CCGCCgccAT
51  CctcgCCGCC TGCCAAAgca gGAGCATCCA AACCTTTCCG CAACCCGACA
101 CATCCGTCAT CAACGGCCCG GACCGGCCGG CCGGCATCCC CGACCCCGCC
151 GGAACGACGG TTGCCGGCGG CGGGGCCGTC TATACCGTTG TGCCGCACCT
201 GTCCATGCCC CACTGGCGCG CGCaggATTT TGCCAAAAGC CTGCAATCCT
251 TCCGCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301 TGCGCCCAAG CCTTTCAAAC CCCCCTGCAT TCCTTTCAGG CAAAGcGgTT
351 TTTTGAACGC TATTTCACGC cgtGGCaggt tgcaggcaAC GGAAGcCTTG
401 Caggtaagggt TACCGGCTAT TACGAACCGG TGCTGAAGGG CGACGGCAGG
451 CGGACGGAAC GGGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCCTGCCG GTTTCGGGG CGGAAAAAAC CTTGTCCGCA
551 TCAGGCAGac ggGAAAAAAC AGCGGCACGA TCGACAATGC CGGCGGCACG
601 CATACGCGCC ACCTCTCCCG ATTCCCATC ACCGCGCGCA CAACGGcaat
651 caaaGGCAGG TTTGAaggAA GCCGCTTCCT CCCTTACCAC ACGCGCAACC
701 AAATcaacGG CGGCGcgctT GACGGCAAag cccCATCTCT CggttacgcC
751 GAagaccCcg tgaacttTT TTTCATGCAC AtccaaggCT CGGGCCGCTT
801 GAAAACCCcg tccggcaaat acatCCGCat cggATacgcc gacAAAAACG
851 AACatccgTa tgtttccatc ggACGctaTA TGGCGGACAA AGGCTACCTC
901 AAGctcgggc agACCTCGAT GCAGGgcata aaagcCTATA TGCGGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC GGCAATGAGG GCCCGTTCGG CGCACTGGGC
1051 ACGCCACTGA TGGGGGAATA CGCCGGCGCA ATCGACCGGC ACTACATTAC
1101 CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CAGGCAGCGC GATCAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGTTAC GGCGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGGATACGT CTGGCAGCTC CTGCCCCAAG
1301 GCATGAAGCC CGAATACCGC CCGTGA

```

This corresponds to the amino acid sequence <SEQ ID 3072; ORF 919.ng>:

```

g919.pep
  1 MKKHLLRSAL YGIAAAAILAA CQSRSIQTFP QPDTSVINGP DRPAGIPDPA
 51 GTTVAGGGAV YTVVPHLSMP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SFQAKRFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDGR
151 RTERARFPIY GIPDDFISVP LPAGLRGGKN LVRIQTGKN SGTIDNAGGT
201 HTADLSRFPI TARTTAIKGR FEGRSFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KAYMRQNPQR LAEVLGQNPS YIFFRELAYS GNEGPVGALG
351 TPLMGEYAGA IDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

ORF 919 shows 95.9 % identity over a 441 aa overlap with a predicted ORF (ORF 919.ng) from *N. gonorrhoeae*:

```

m919/g919

      10      20      30      40      50      60
m919.pep MKKYLFR AALYGI AAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
          |||::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g919      MKKHLLRSALYGI AAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
          10      20      30      40      50      60

      70      80      90     100     110     120
m919.pep YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHVSFQAKQFFER
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g919      YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHVSFQAKRFFER
          70      80      90     100     110     120

      130     140     150     160     170     180
m919.pep YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g919      YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTERARFPIYGIIPDDFISVPLPAGLRGGKN
          130     140     150     160     170     180

      190     200     210     220     230     240
m919.pep LVRIQTGKNSGTIDNTGGTHTADLSRFPI TARTTAIKGRFEGRSFLPYHTRNQINGGAL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g919      LVRIQTGKNSGTIDNAGGHTADLSRFPI TARTTAIKGRFEGRSFLPYHTRNQINGGAL
          190     200     210     220     230     240

      250     260     270     280     290     300
m919.pep DGKAPILGYAEDPVELFFMH IQGSGRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g919      DGKAPILGYAEDPVELFFMH IQGSGRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL
          250     260     270     280     290     300

      310     320     330     340     350     360
m919.pep KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYI FFRELAGSSNDGPVGALGTPLMGEYAGA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g919      KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYI FFRELAGSGNEGPVGALGTPLMGEYAGA
          310     320     330     340     350     360

      370     380     390     400     410     420
m919.pep VDRHYITLGAPL FVATAHPVTRKALNRLIM AQDTGSAIKGAVRVDYFWGYGDEAGELAGK
          :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g919      IDRHYITLGAPL FVATAHPVTRKALNRLIM AQDTGSAIKGAVRVDYFWGYGDEAGELAGK
          370     380     390     400     410     420

```

```

          430      440
m919.pep  QKTTGYVWQLLPNGMKPEYRFX
          |||||
g919      QKTTGYVWQLLPNGMKPEYRFX
          430      440

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3073>:

```

a919.seq
1  ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TGCGGCATCG CCGCCGCCAT
51  CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA
101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC
151 GGAACGACGG TCGGCGGCCG CGGGGCCGTT TATACCGTTG TGCCGCACCT
201 GTCCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAGC CTGCAATCCT
251 TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301 TGCGCCCAAG CCTTTCAAAC CCCCCTCCAT TCCGTTCCAG CAAACAGTT
351 TTTTGAACGC TATTTACGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG
401 CCGGTACGGT TACCGGCTAT TACGAGCCGG TGCTGAAGGG CGACGACAGG
451 CGGACGGCAC AAGCCCGCTT CCGGATTTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTTGTCCGCA
551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
601 CATACCGCCG ACCTCTCCCA ATTCCCCATC ACTGCGCGCA CAACGGCAAT
651 CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC
701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCATACT CGGTTACGCC
751 GAAGACCCCG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
851 AACATCCCTA CGTTTCCATC GGACGCTATA TGGCGGACAA AGGCTACCTC
901 AAGCTCGGGC AGACCTCGAT GCAGGGCATC AAAGCCTATA TGCAGCAAAA
951 CCCGCAACGC CTCGCCGAAG TTTTGGGGCA AAACCCAGC TATATCTTTT
1001 TCCGAGAGCT TACCGGAAGC AGCAATGACG GCCCTGTGGC CGCACTGGGC
1051 ACGCCGCTGA TGGGCGAGTA CGCCGGCGCA GTCGACCGGC ACTACATTAC
1101 CTTGGGCGCG CCCTTATTG TCGCCACCGC CCATCCGGT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTT CTGCCAACG
1301 GTATGAAGCC CGAATACCG CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3074; ORF 919.a>:

```

a919.pep
1  MKKYLFRAL CGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
51  GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SVQAKOFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDGR
151 RTAQAARFPIY GIPDDFISVP LPAGLRSGKA LVRIROTKGN SGTIDNTGGT
201 HTADLSQFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KAYMQONPQR LAEVLGONPS YIFFRELTGS SNDGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

m919/a919 ORFs 919 and 919.a showed a 98.6% identity in 441 aa overlap

```

          10      20      30      40      50      60
m919.pep  MKKYLFRALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
          |||||
a919      MKKYLFRALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
          10      20      30      40      50      60

          70      80      90      100     110     120
m919.pep  YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKOFFER
          |||||
a919      YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKOFFER
          70      80      90      100     110     120

```

m919.pep	130	140	150	160	170	180
	YFTPWQVAGNGSLAGTVTGYE	PVLKGD	DRRTAQARFPIYGIPDDFISVPLPAGLRSGKA			
a919	130	140	150	160	170	180
	YFTPWQVAGNGSLAGTVTGYE	PVLKGD	DRRTAQARFPIYGIPDDFISVPLPAGLRSGKA			
m919.pep	190	200	210	220	230	240
	LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL					
a919	190	200	210	220	230	240
	LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL					
m919.pep	250	260	270	280	290	300
	DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL					
a919	250	260	270	280	290	300
	DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL					
m919.pep	310	320	330	340	350	360
	KLQOTSMQGIKSYMRQNPORLAEVLGQNPSYIFFREL	AGSSNDGPVGALGTPLMGEYAGA				
a919	310	320	330	340	350	360
	KLQOTSMQGIKAYMQQNPORLAEVLGQNPSYIFFREL	TGSSNDGPVGALGTPLMGEYAGA				
m919.pep	370	380	390	400	410	420
	VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK					
a919	370	380	390	400	410	420
	VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK					
m919.pep	430	440				
	QKTTGYVWQLLPNGMKPEYRPX					
a919	430	440				
	QKTTGYVWQLLPNGMKPEYRPX					

121 and 121-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3075>:

m121.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCC TACCCCGGCA GGTACGCCG CCAATTGCTG
151 GATTTGCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAACCTC CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATAACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
401 XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
451 XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
501 XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
551 XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
601 XXXXXXCAGC TTCCTTACGA CAAAACGGT GCAAAGTCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACGCCACCC TAAAGACACG GGGCGCGAAC TGTTCGCCAT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGTTT ACCGCGCAAA CCGTTTGCGA CGCGCTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGACG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGnATTG
1001 CGTGGTTGGC GCGGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA

```

1051 GCAACCGGCG CATCCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG  
1101 A

This corresponds to the amino acid sequence <SEQ ID 3076; ORF 121>:

m121.pep

```

1  METQLYIGIM SGTSMGDGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRQLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL Axxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
151 xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
201 xxQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251 ETYLDGGENR YDVLRTLSRF TAQTVCDAYS HAAADARQMY ICDGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3077>:

g121.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
51  GCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCTGACC GGTTGCGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCGGAA CTGCTGTGCA
251 GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATAACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACTGa cgcggattttT TACCGTCggc gacttcCGCA
401 GCCGCGACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCCgcCTTT
451 CACGAAGCCC TGTTCCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCggCGCA CCCCgcCTTCG
551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
601 cactGGcagc TGCCTTACGA CAAAaAcggt gcAAAGgcgg cacAAGGCAA
651 catatTGcCg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCAC
701 AACCCcacc aaAAAGCACG GGgcGCGaac TgtttgcccT AAattggctc
751 gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgt
801 ttcccgattc accgcgcaaa ccgTttggga cgccgtctca CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTtg
1001 cgtggttgC GCGGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCGACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3078; ORF 121.ng>:

g121.pep

```

1  METQLYIGIM SGTSMGDGADA VLVRMDGGKW LGAEGHAFTP YPDRLLRRKLL
51  DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPFAF
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVWDAYS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from *N. gonorrhoeae*:

m121/g121

```

          10          20          30          40          50          60
m121.pep  METQLYIGIMSGTSMGDGADAVLIRMDGGKWLGAEGHAFTYPYGRLLRRQLLDLQDTGADEL
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g121       METQLYIGIMSGTSMGDGADAVLVRMDGGKWLGAEGHAFTYPYDRLLRRKLLDLQDTGTDEL
          10          20          30          40          50          60
          70          80          90         100         110         120
m121.pep  HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEGYSIQLADLPLL
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

g121      HRSRMLSQELSRLYAQTAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
           70      80      90      100      110      120
           130      140      150      160      170      180
m121.pep  AXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
           | : : : : :
g121      AELTRIFTVGDFRSRDLAAGGQGAPLVPAPFHEALFRDDRETRVVLNIGGIANISVLPPGA
           130      140      150      160      170      180
           190      200      210      220      230      240
m121.pep  XXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST
           : : : |||||:|||||:|||||:|||||:|||||
g121      PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
           190      200      210      220      230      240
           250      260      270      280      290      300
m121.pep  GRELFALNWLETYLDGGENRYDVLRTLRSRFTAQTVCDVASHAAADARQMYICDGGIRNPV
           |||||:|||||:|||||:|||||:|||||:|||||
g121      GRELFALNWLETYLDGGENRYDVLRTLRSRFTAQTVWDVASHAAADARQMYICGGGIRNPV
           250      260      270      280      290      300
           310      320      330      340      350      360
m121.pep  LMADLAECFGTRVSLHSTADLNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
           |||||:|||||:|||||:|||||:|||||:|||||
g121      LMADLAECFGTRVSLHSTAE LNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
           310      320      330      340      350      360

m121.pep  XAGYYYY
           |||||
g121      GAGYYYY

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3079>:

```

a121.seq
1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAAGTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCACGCA CCCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCAC
701 AACCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCATTAC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGCGC CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3080; ORF 121.a>:

```

a121.pep
1  METQLYIGIM SGTSMGDGADA VLIRMDGGKW LGAEGHAFTP YPGLRRLKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAEL LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAP
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLRSR TAQTVFDAVS HAAADARQMY ICGGGIRNPV

```

301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AFAWMAACW VNRIPGSPHK  
 351 ATGASKPCIL GAGYYY\*

m121/a121 ORFs 121 and 121.a showed a 74.0% identity in 366 aa overlap

m121.pep	10	20	30	40	50	60
	METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPPYGRRLRRQLLDLQDTGADEL					
a121	METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPPYGRRLRRQLLDLQDTGADEL					
	10	20	30	40	50	60
m121.pep	70	80	90	100	110	120
	HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPENHYSIQLADLPLL					
a121	HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPENHYSIQLADLPLL					
	70	80	90	100	110	120
m121.pep	130	140	150	160	170	180
	AXXX					
a121	AERTQIFTVGDFRSRDLAGGQGAPLVPFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
m121.pep	190	200	210	220	230	240
	XXXXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST					
a121	PAFGFDTGPGNMLMDAWMOAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
m121.pep	250	260	270	280	290	300
	GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDVASHAADARQMYICDGGIRNPV					
a121	GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVFVASHAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
m121.pep	310	320	330	340	350	360
	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGTRVSLHSTAE LNLDPQWVEA AFAWMAACWVNRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYYX					
a121	GAGYYYYX					

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3081>:

m121-1.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CCAATTGCTG
151 GATTTGCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCACGCTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCGCAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CGCGGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TCGCGCCGGC GGACAAGGCG CGCCACTCGT CCCC GCCTTT
451 CACGAAGCCC TGTTCGCGA CAACAGGGAA ACACGCGCG TACTGAACAT
501 CCGCGGGATT GCCAACATCA GCGTACTCCC CCCC GACGCA CCGCCTTCG
551 GCTTCGACAC AGGCGCGGGC AATATGCTGA TGGACGCGTG GACGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACCCACCC TAAAAGCAGG GGGCGCGAAC TGTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGTTTT ACCGCGCAAA CCGTTTGC GA CGCGTCTCA CACGCAGCGG

```



```

851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGNATTG
1001 CGTGGTTGGC GCGGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG ANCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3082; ORF 121-1>:

m121-1.pep

```

1 METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRQLL
51 DLQDTGADEL HRSRILSOEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL AERTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDNRE TRAVLNIGGI ANISVLPPDA PAFGFDTPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVCDAYS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

m121-1/g121 ORFs 121-1 and 121.ng showed a 95.6% identity in 366 aa overlap

	10	20	30	40	50	60
m121-1.pep	METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLLRQLLDLQDTGADEL					
	:     :     :     :     :					
g121	METQLYIGIMSGTSMDGADAVLVRMDGGKWLGAEGHAFTPYPDRLLRKLDDLQDTGTDEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121-1.pep	HRSRILSOELSRLYAQTAAELLCSONLAPSDITALGCHGQTVRHAPHEGYISQLADLPLL					
	:     :     :     :     :					
g121	HRSRMLSOELSRLYAQTAAELLCSONLAPCDITALGCHGQTVRHAPHEGYISQLADLPLL					
	70	80	90	100	110	120
	130	140	150	160	170	180
m121-1.pep	AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA					
	:     :     :     :     :					
g121	AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121-1.pep	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	:     :     :     :     :					
g121	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121-1.pep	GRELFALNWLETYLDGGENRYDVLRTLSRF TAQTVCDAYS HAAADARQMYICGGGIRNPV					
	:     :     :     :     :					
g121	GRELFALNWLETYLDGGENRYDVLRTLSRF TAQTVWDAYS HAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m121-1.pep	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
	:     :     :     :     :					
g121	LMADLAECFGTRVSLHSTAE LNLDPQWVEA AAFWLAACWINRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121-1.pep	XAGYYYYX					
g121	GAGYYYYX					

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3083>:

a121-1.seq

```

1 ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG

```

```
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAAGTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TCGCGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCAGCGCA CCCCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACCCACCC TAAAAGCACG GGGCGCGAAC TGTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TCGCGACGCT
801 TTCCCGATT ACCGCGCAA CCGTTTTTCA CGCCGTCTCA CACGACGCG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGGAT ATTATTATTG
1101 A
```

This corresponds to the amino acid sequence <SEQ ID 3084; ORF 121-1.a>:

a121-1.pep

```
1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRLKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAAPS DITALGCHGQ
101 TVRHAPESY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMDLAECFG TRVSLHSTAE LNLDPQWVEA AFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*
```

m121-1/a121-1 ORFs 121-1 and 121-1.a showed a 96.4% identity in 366 aa overlap

m121-1.pep	10	20	30	40	50	60
	METQLYIGIMSGTSMGDADAVLIRMDGGKW	LGAEGHAFTPYPGRLLRRLKLLDLQDTGADEL				
a121-1	10	20	30	40	50	60
	METQLYIGIMSGTSMGDADAVLIRMDGGKW	LGAEGHAFTPYPGRLLRRLKLLDLQDTGADEL				
m121-1.pep	70	80	90	100	110	120
	HRSRILSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAP	EHGYSIQADLPLL				
a121-1	70	80	90	100	110	120
	HRSRMLSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAP	ESYSVQLADLPLL				
m121-1.pep	130	140	150	160	170	180
	AERTQIFTVGDFRSRDLAAGGQGAPLVPFHEALFRDNRETRAVLNIGGI	ANISVLPPDA				
a121-1	130	140	150	160	170	180
	AERTQIFTVGDFRSRDLAAGGQGAPLVPFHEALFRDDRETRAVLNIGGI	ANISVLPPDA				
m121-1.pep	190	200	210	220	230	240
	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHP	YFAQPHPKST				
a121-1	190	200	210	220	230	240
	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHP	YFAQPHPKST				
m121-1.pep	250	260	270	280	290	300
	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMY	ICGGGIRNPV				
a121-1	250	260	270	280	290	300
	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFDAVSHAAADARQMY	ICGGGIRNPV				

	310	320	330	340	350	360
m121-1.pep	LMADLAECFGTRVSLHSTADLNLD	PQWVEAAXFAWLAACWINRIPG	SPHKATGASKPCIL			
a121	LMADLAECFGTRVSLHSTAE	NLDLPQWVEAAAFAWMAACWVNRI	PGSPHKATGASKPCIL			
	310	320	330	340	350	360

m121-1.pep	XAGYYYYX
a121	GAGYYYYX

## 128 and 128-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3085>:

```

m128.seq (partial)
1  ATGACTGACA ACGCACTGCT CCATTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACTCA ACTGCGTCGC CGACACGCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTTCGAC ACCCTCTCCC CCGCACAAA AACCAAACTC AACCAC
1  TACGCCAGCG AAAAACTGCG CGAAGCCAAA TACGCGTTCA GCGAAACCGA
51  wGTCAAAAAA TAYTCCCyG TCGGCAAwGT ATTAAACGGA CTGTTTCGCCC
101 AAmTCAAAAA ACTmTACGGC ATCGGATTTA CCGAAAAAAC yGTCCCCGTC
151 TGGCACAAG ACGTGCCTA TtkTGAATTG CAACAAAACG GCGAAmCCAT
201 AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
251 CGTGGATGAA CGACTACAAA GGCCGCCGCC GTTTTTCAGA CGGCACGCTG
301 CAAyTGCCCA CCGCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGGCGG
351 CAGGGAAGCC CGCyTGAGCC ACGACGAAAT CCTCATCCTC TTCCACGAAA
401 CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
451 TCCGGCATCA ACGGCGTaKa ATGGGACGCG GTCGAACTGC CCAGCCAGTT
501 TATGGAATAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
551 ACGAAGAAAC CGGcgTTCCC yTGCCGAAAG AACTCTTsGA CAAAwTGCTC
601 GCCGCCAAAA ACTTCCAAsG CGGCATGTTT yTsGTCCGGC AAwTGGAGTT
651 CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
701 AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTCATC
751 CAGCCGCCCG AATACAACCG CTTGCCTTG AGCTTCGGCC ACATCTTCGC
801 AGGCGGCTAT TCCGCAGCTn ATTACAGCTA CGCGTGGGCG GAAGTATTGA
851 GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
901 GGCAAAACGCT TTTGGCAGGA AATCCTCGCC GTCGGGGnAT CGCGCAGCGG
951 nGCAGAAATCC TTCAAAGCCT TCCGCGGCCG CGAACCAGAGC ATAGACGCAC
1001 TCTTGCGCCA CAGCGGTTTC GACAACGCGG TCTGA

```

This corresponds to the amino acid sequence <SEQ ID 3086; ORF 128>:

```

m128.pep (partial)
1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWVA
51  NTVEPLTGIT ERVGRWGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NH

//

1  YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGFTTEKTVPV
51  WHKDVRYXEL QONGEXIGGV YMDLYAREGK RGGAWMNDYK GRRRFSDBGTL
101 QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
151 SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
201 AAKNFQXGMF XVRQXEFALF DMMIYSEDDE GRLKNWQQVL DSVRKKVAVI
251 QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT
301 GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHSGF DNAV*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3087>:

```

g128.seq
1  atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTTaatca
51  aatccaaacc gaagACAtca AACCCGCCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTGT
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCG GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCAGGACC
401 TGCGCGATT TCGTATTGAGC GGC CGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGccaaa cTGCTCGGCT TTAAAAATTA CGCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCCGCCGC GAACACCTCG GTCTCGCCGA CCCGAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CCGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACCGGCGCG
1251 CGCGTGGATG AACGACtaca AAGGCCGCCG CCGCTTTGCC GACGgcacGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AacCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAACggcgta GAATGGGACG CGGTCGAACT GCCCAGCCAG
1501 TTTATGAAA ACTTCGTTTG GGAATACAAT GTATTGCAC AAATGTCCGC
1551 CCACGAAGAA AccgGCGAGC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TcgCGCCAA AAATTTCCAG CGCGTATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TCGATATGAT GATTTACAGT GAAAGCGACG AATGCCGTCT
1701 GAAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751 TCCAACGCC CGAATAACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801 GCcggcGGCT ATTCCGCAGG CTATTACAGC TACGCATGGG CCGAAGTCct
1851 cAGCACCGAT GCCTACGCCG CCTTTGAAGA AAGcGACGac gtcGCCGCCA
1901 CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
1951 gcgCGGGAAT CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTGCTGCGC CAaagcggtT TCGACAACGC gGcttga

```

This corresponds to the amino acid sequence <SEQ ID 3088; ORF 128.ng>:

```

g128.pep
1  MIDNALLHLG EEPRFNQIQT EDIKPAVQTA IAEARGQIAA VKAQTHTGWA
51  NTVERLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDSL YAGEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVVHKDVR YFELQONGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTQLQLPTAY LVCNFAPFVG
451 GKEARLSHDE ILTLFHETGH GLHLLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS ESDECRLKNW QQVLDVRKE VAVIQPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSTD AYAAFEESDD VAATGKRFWQ EILAVGGSRS

```

651 AAESFKAFRG REPSIDALLR QSGFDNAA\*

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng)  
from *N. gonorrhoeae*:

m128/g128

```

      10      20      30      40      50      60
g128.pep  MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQTHGTWANTVERLTGIT
          | | | | | | | | | | : | | | | | | | | | | : | | | | | | | | | |
m128      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHGTWANTVEPLTGIT
          10      20      30      40      50      60

      70      80      90     100     110     120
g128.pep  ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
          | | | | | | | | | | : | | | | | | | | | | : | | | | | | | | | |
m128      ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90     100     110     120

      130     140     150     160     170     180
g128.pep  TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
          | | | | | | | | | | : | | | | | | | | | | : | | | | | | | | | |
m128      TLSPAQKTKLNH
          130
          //

                                340     350     360
g128.pep                                YAGEKLRKAYAFSETEVKKYFPVGKVLG
                                | | : | | | | | | | | | | | | | | | | | |
m128                                YASEKLRKAYAFSETXVKKYFPVGXVLNG
                                10      20      30

      370     380     390     400     410     420
g128.pep  LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK
          | | | | | | | | | | : | | | | | | | | | | : | | | | | | | | | |
m128      LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWMNDYK
          40      50      60      70      80      90

      430     440     450     460     470     480
g128.pep  GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVDELGV
          | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128      GRRRFSGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVDELGV
          100     110     120     130     140     150

      490     500     510     520     530     540
g128.pep  SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQXGMF
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128      SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQXGMF
          160     170     180     190     200     210

      550     560     570     580     590     600
g128.pep  LVRQMEFALFDMMIYSESDCRLKNWQQVLDVSRKEVAVIQPPEYNRFANSFGHIFAGGY
          | | | | | | | | | | : | | | | | | | | | | : | | | | | | | | | |
m128      XVRQXEFALFDMMIYSEDDGRLKNWQQVLDVSRKKVAVIQPPEYNRFALSFGHIFAGGY
          220     230     240     250     260     270

      610     620     630     640     650     660
g128.pep  SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS
          | | : | | | | | | | | : | | | | | | | | | | : | | | | | | | | | |
m128      SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRGREPS
          280     290     300     310     320     330
```

```

          670      679
g128.pep  IDALLRQSGFDNAAX
          |||||:|||||:
m128      IDALLRHSGFDNAVX
          340

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3089>:

```

a128.seq
1  ATGACTGACA  ACGCACTGCT  CCATTTGGGC  GAAGAACCCC  GTTTTGATCA
51  AATCAAAACC  GAAGACATCA  AACCCGCCCT  GCAAACCGCC  ATTGCCGAAG
101 CGCGCGAACA  AATCGCCGCC  ATCAAAGCCC  AAACGCACAC  CGGCTGGGCA
151 AACACTGTCT  AACCCTGAC   CGGCATCACC  GAACGCGTCG  GCAGGATTTG
201 GGGCGTGGTG  TCGCACCTCA  ACTCCGTAC  CGACACGCCC  GAACTGCGCG
251 CCGCCTACAA  TGAATTAATG  CCCGAAATTA  CCGTCTTCTT  CACCGAAATC
301 GGACAAGACA  TCGAGCTGTA  CAACCGCTTC  AAAACCATCA  AAAACTCCCC
351 CGAGTTCGAC  ACCCTCTCCC  ACGCGCAAAA  AACCAAACTC  AACCACGATC
401 TGCGCGATT  CGTCCTCAGC  GCGCGGGAAC  TGCCGCCCGA  ACAGCAGGCA
451 GAATTGGCAA  AACTGCAAAC  CGAAGGCGCG  CAACTTTCCG  CCAAATCTCT
501 CCAAAACGTC  CTAGACGCGA  CCGACGCGTT  CGGCATTTAC  TTTGACGATG
551 CCGCACCGCT  TGCCGGCATT  CCCGAAGACG  CGCTCGCCAT  GTTTGCCGCT
601 GCCGCGCAA  GCGAAGGCAA  AACAGGCTAC  AAAATCGGTT  TGCAGATTCC
651 GCACTACCTC  GCCGTCATCC  AATACGCCGA  CAACCGCAAA  CTGCGCGAAC
701 AAATCTACCG  CGCCTACGTT  ACCCGCGCCA  GCGAGCTTTC  AGACGACGGC
751 AAATTCGACA  ACACCGCCAA  CATCGACCGC  ACGCTCGAAA  ACGCCCTGCA
801 AACCGCCAAA  CTGCTCGGCT  TCAAAAATA  CGCCGAATTG  TCGCTGGCAA
851 CCAAATGGC  GGACACCCCC  GAACAAGTTT  TAAACTTCCT  GCACGACCTC
901 GCCCGCGCG  CCAAACCTTA  CGCCGAAAA  GACCTCGCCG  AAGTCAAAGC
951 CTTGCCCCG  GAAAGCCTCG  GCCTCGCCGA  TTTGCAACCG  TGGGACTTGG
1001 GCTACGCCG  CGAAAACTG  CGCGAAGCCA  AATACGCATT  CAGCGAAACC
1051 GAAGTCAAAA  AATACTTCCC  CGTCGGCAA  GTATTAAACG  GACTGTTCCG
1101 CCAAATCAA  AAATCTACG  GCATCGGATT  TACCGAAAA  ACCGTCCCCG
1151 TCTGGCACA  AGACGTGCGC  TATTTTGAAT  TGCAACAAA  CGGCGAAACC
1201 ATAGGCGCG  TTTATATGGA  TTTGTACGCA  CGCGAAGGCA  AACGCGGCGG
1251 CGCGTGGATG  AAGGCTGACA  AAGGCGCGCG  CCGTTTTTCA  GACGGCACGC
1301 TGCAACTGCC  CACCGCCTAC  CTCGTCTGCA  ACTTCACCCC  GCCCGTCGGC
1351 GGCAAAGAAG  CCCGCTTGAG  CCATGACGAA  ATCCTCACC  TCTTCCACGA
1401 AACC GGACAC  GGCCTGCACC  ACCTGCTTAC  CCAAGTCGAC  GAACTGGGCG
1451 TATCCGGCAT  CAACGGCGTA  GAATGGGACG  CAGTCGAACT  GCCCAGTCAG
1501 TTTATGAAA  ATTCGTTTG  GGAATACAAT  GTCTTGCGCG  AAATGTCCGC
1551 CCACGAAGAA  ACCGGCGTTC  CCTGCGCGAA  AGAACTCTTC  GACAAAATGC
1601 TCGCGCCAA  AAATCTCAA  CGCGGAATGT  TCCTCGTCCG  CCAAATGGAG
1651 TTCGCCCTCT  TTGATATGAT  GATTACAGC  GAAGACGACG  AAGGCCGTCT
1701 GAAAACTGG  CAACAGGTTT  TAGACAGCGT  GCGCAAAGAA  GTCGCCGTCT
1751 TCCGACCGCC  CGAATACAAC  CGCTTCGCCA  ACAGCTTCGG  CCACATCTTC
1801 GCAGGCGGCT  ATTCCGAGG  CTATTACAGC  TACGCGTGGG  CGGAAGTATT
1851 GAGCGCGGAC  GCATACGCCG  CCTTTGAAGA  AAGCGACGAT  GTCGCCGCCA
1901 CAGGCAAACG  CTTTGGCAG  GAAATCCTCG  CCGTCGGCGG  ATCGCGCAGC
1951 GCGGCAGAA  CTTCAAAGC  CTTCCGCGGA  CGCGAACCGA  GCATAGACGC
2001 ACTCTTGCGC  CACAGCGGCT  TCGACAACGC  GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3090; ORF 128.a>:

```

a128.pep
1  MTDNALLHLG  EEPFRDQIKT  EDIKPALQTA  IAEAREQIAA  IKAQTHTGWA
51  NTVEPLTGIT  ERVGRWGVV  SHLNSVTDTP  ELRAAYNELM  PEITVFFTEI
101 GQDIELYNRF  KTIKNSPEFD  TLSHAQKTKL  NHDLRDFVLS  GAELPPEQQA
151 ELAKLQTEGA  QLSAKFSQNV  LDATDAFGIY  FDDAAPLAGI  PEDALAMFAA
201 AAQSEGKTGY  KIGLQIPHYL  AVIQYADNRK  LREQIYRAYV  TRASELSDDG
251 KFDNTANIDR  TLENALQTAK  LLGFKNYAEL  SLATKMADTP  EQVLNLFHDL
301 ARRAKPYAEK  DLAEVKAFAR  ESLGLADLQP  WDLGYAGEKL  REAKYAFSET
351 EVKKYFPVGK  VLNGLFAQIK  KLYGIGFTEK  TVPVWHKDVR  YFELQQNGET
401 IGGVYMDLYA  REGKRGGAWM  NDYKGRRRFS  DGTLLQLPTAY  LVCNFTPPVG
451 GKEARLSHDE  ILTLFHETGH  GLHLLTQVD  ELGVSGINGV  EWDVAVELPSQ
501 FMENFVWEYN  VLAQMSAHEE  TGVPLPKELF  DKMLAAKNFQ  RGMFLVRQME

```

551 FALFDMMIYS EDDEGR LKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF  
601 AGGYSAGYYS YAWAEVLSAD AYA AFEESDD VAATGKRWFQ EILAVGGSR S  
651 AAESFKA FRG REPSIDALLR HSGFDNAA\*

m128/a128 ORFs 128 and 128.a showed a 66.0% identity in 677 aa overlap

m128.pep	10	20	30	40	50	60
	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT					
a128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT					
	10	20	30	40	50	60
m128.pep	70	80	90	100	110	120
	ERVGRIGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
a128	ERVGRIGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
m128.pep	130	TLSPAQKTKLNH-----				
a128	130	140	150	160	170	180
	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
m128.pep	-----					
a128	190	200	210	220	230	240
	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLRQIYRAYV					
m128.pep	-----					
a128	250	260	270	280	290	300
	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
m128.pep	-----					
	140 150					
	-----YASEKLREAKYAFSETXVKKYFPVGX					
a128	310	320	330	340	350	360
	ARRAKPYAEKDLAEVKAFARESGLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGK					
m128.pep	160	170	180	190	200	210
	VLNGLFAQXKKLYGIGFTEKTPVWHKDVRYXELQONGEXIGGVYMDLYAREGKRGGAWM					
a128	370	380	390	400	410	420
	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQONGETIGGVYMDLYAREGKRGGAWM					
m128.pep	220	230	240	250	260	270
	NDYKGRRRFSDGTLQLPTAYLVCNFPVGGREARLSHDEILILFHETGHGLHLLTQVD					
a128	430	440	450	460	470	480
	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
m128.pep	280	290	300	310	320	330
	ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ					
a128	490	500	510	520	530	540
	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
m128.pep	340	350	360	370	380	390
	XGMFXVRQXEFALFDMMIYSEDDEGR LKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF					
a128	550	560	570	580	590	600
	RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF					

m128-1.seq

```
1 ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51 AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTGC AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACTCA ACTCGTCGC CGACACGCCG GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTCTT CACCGAAAT
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTTCGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACACGATC
401 TGCGCGATTT CGTCTCAGC GGC GCGGAAC TGCCGCGCGA ACAGCAGGCA
451 GAAGCTGGCAA AACTGCAAA CGAAGCGCG CAACTTTCCG CCAAATTCCT
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAT GTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT TTTGCGGCC
601 GCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
651 ACACTACTCT GCCGTATCC AATACGCCCA CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCTACGTT ACCCGGCCA GCGAACTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCG ACGTCGCAA ACGCCTGCA
801 AACC GCCAAA CTGCTCGGCT TCAAAAATA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGC CCAAACCTTA CGCCGAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAAACTG CCGCGAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTATTAAACG GACTGTTGCG
1101 CCAAATCAA AAACCTACG GCATCGGATT TACCGAAAA ACCGTCCCCG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAA CGGCGAAACC
1201 ATAGGCGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACCGCGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCG CGCTTTTCA GACGCGACG
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGAAG CCGCCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
1401 AACCGGACG GGGCTGCAC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TATCGGCAT CAACGGCGTA GAATGGGAC CGGTGCAAT GCCCAGCCAG
1501 TTTATGGAAT ATTTGTTTG GGAATACAA GTCTTGGCAC AAATGTGAG
1551 CCACGAAGAA ACCGCGCTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAATTTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
1651 TTGCGCCTCT TTGATATGAT GATTTACAG GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAA GTCACCGTCA
1751 TCCAGCGGCC CGAATACAAC CGCTTCGCTT TGAGTCTCGG CCACATCTTC
1801 GACGCGGCT ATTCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCC CTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGCGAAAAC CTTTGGCAG GAAATCCTCG CCGTCGCGG ATCGCGCAGC
1951 CGGCAGAAAT CTTCAAAGC GTTCGCGCG CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGTT TCGACAACGC GGTCTGA
```

ml28-1.pcp.  
1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHGWA  
51 NTVEPLTGIT ERVGRIWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI



```

101  GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201  AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251  KFDNTANIDR TLANALQTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301  ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351  EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFAPPVG
451  GREARLSHDE ILILFHETGH GLHLLLTQVD ELGVSGINGV EWDDELPSQ
501  FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551  FALFDMMIYS EDDEGRLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAV*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3093>:

g128-1.seq (partial)

```

1   ATGATTGACA ACGCACTGCT CCACTTGGGC GAAGAACCCC GTTTTAATCA
51  AATCAAAACC GAAGACATCA AACCCGCCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCGGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
401 TGCGCGATTT CGTATTGAGC GCGCGGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAA CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 CCGCGCAGAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGCCAAA CTGCTCGGCT TTAATAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACGCCC GAACAGGTTT TAACTTCCT GCACGACCTC
901 CCGCGCCGCG CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCCG GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTTCG
1101 CCAAATCAAA AACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGCGCG
1251 CGCGTGGATG AACGACTACA AAGCCGCGCG CCGCTTTGCC GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AACC GGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAACGGCGTA AAA

```

This corresponds to the amino acid sequence <SEQ ID 3094; ORF 128-1.ng>:

g128-1.pep (partial)

```

1   MIDNALLHLG EEPRFNQIKT EDIKPAVQTA IAEARGQIAA VKAQHTGTGA
51  NTVERLTGIT ERVGRWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEKGTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLGYASEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQONGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLLTQVD ELGVSGINGV K

```

m128-1/g128-1 ORFs 128-1 and 128-1.ng showed a 94.5% identity in 491 aa overlap

10

20

30

40

50

60

g128-1.pep	MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQHTGTGWANTVERLTGIT
m128-1	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGTGWANTVEPLTGIT
	10 20 30 40 50 60
g128-1.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
m128-1	ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
	70 80 90 100 110 120
g128-1.pep	TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSONVLDATDAFGIY
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSONVLDATDAFGIY
	130 140 150 160 170 180
g128-1.pep	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
	190 200 210 220 230 240
g128-1.pep	TRASELSNDGKFDNTANIDRTLLENALKTAKLLGFKNYAELSLATKMADTPEQVNLFLHDL
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAELGFKNYAELSLATKMADTPEQVNLFLHDL
	250 260 270 280 290 300
g128-1.pep	ARRAKPYAEKDLAEVKAFAREHLGLADPQPWDLSYAGEKLREAKYAFSETEVKKYFPVGK
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLPWDLYASEKLREAKYAFSETEVKKYFPVGK
	310 320 330 340 350 360
g128-1.pep	VLGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQONGKTIGGVYMDLYAREGKRGGAWM
m128-1	VLNLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQONGKTIGGVYMDLYAREGKRGGAWM
	370 380 390 400 410 420
g128-1.pep	NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHLHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHLHLLTQVD
	430 440 450 460 470 480
g128-1.pep	ELGVSGINGVK
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
	490 500 510 520 530 540

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3095>:

a128-1.seq

1	ATGACTGACA	ACGCACTGCT	CCATTGGGC	GAAGAACCCC	GTTTTGATCA
51	AATCAAAACC	GAAGACATCA	AACCCGCCCT	GCAAACCGCC	ATTGCCGAAG
101	CGCGCGAACA	AATCGCCGCC	ATCAAAGCCC	AAACGCACAC	CGGCTGGGCA
151	AACACTGTCTG	AACCCCTGAC	CGGCATCACC	GAACGCGTCG	GCAGGATTGT
201	GGGCGTGGTG	TCGCACCTCA	ACTCCGTCAC	CGACACGCCC	GAACGCGCG
251	CCGCCTACAA	TGAATTAATG	CCCGAAATTA	CCGTCTTCTT	CACCGAAATC
301	GGACAAGACA	TCGAGCTGTA	CAACCGCTTC	AAAACCATCA	AAAACCTCCC
351	CGAGTTCGAC	ACCCTCTCCC	ACGCGCAAAA	AACCAAATC	AACCACGATC
401	TGCGCGATTT	CGTCCTCAGC	GGCGCGGAAC	TGCCGCGGCA	ACAGCAGGCA
451	GAATTGGCAA	AACTGCAAC	CGAAGGCGCG	CAACTTCCG	CCAAATTCTC

```

501 CAAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCGCGCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCCCCG GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGCGGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCGC CCGTTTTTCA GACGGCACCG
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAAC TCCAGTCAG
1501 TTTATGGAAT ATTTGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAATCTCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCT
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCGCCCA
1901 CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACC GAATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3096; ORF 128-1.a>:

a128-1.pep

```

1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGTGWA
51 NTVEPLTGIT ERVGRWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTTY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTA LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVVHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDVAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAA*

```

m128-1/a128-1 ORFs 128-1 and 128-1.a showed a 97.8% identity in 677 aa overlap

```

          10      20      30      40      50      60
a128-1.pep MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGTGWANTVEPLTGIT
          |||
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGTGWANTVEPLTGIT
          10      20      30      40      50      60

          70      80      90      100     110     120
a128-1.pep ERVGRWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          |||
m128-1      ERVGRWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90      100     110     120

          130     140     150     160     170     180

```

a128-1.pep	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAE LAKLQTEGAQLSAKFSONVLDATDAFGIY
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAE LAKLQTEGAQLSAKFSONVLDATDAFGIY
	130 140 150 160 170 180
a128-1.pep	190 200 210 220 230 240
	FDDAAPLAGIPEDALAMFAAAAQSEKGTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAAQSEKGTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
	190 200 210 220 230 240
a128-1.pep	250 260 270 280 290 300
	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVNLFLHDL
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVNLFLHDL
	250 260 270 280 290 300
a128-1.pep	310 320 330 340 350 360
	ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLG YAGEKLREAKYAFSETEVKKYFPVGK
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLQPWDLG YASEKLREAKYAFSETEVKKYFPVGK
	310 320 330 340 350 360
a128-1.pep	370 380 390 400 410 420
	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQONGETIGGVYMDLYAREGKRGGAWM
m128-1	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQONGETIGGVYMDLYAREGKRGGAWM
	370 380 390 400 410 420
a128-1.pep	430 440 450 460 470 480
	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD
	430 440 450 460 470 480
a128-1.pep	490 500 510 520 530 540
	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
	490 500 510 520 530 540
a128-1.pep	550 560 570 580 590 600
	RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLD SVRKEVAVVRPPEYNRFANSFGHIF
m128-1	RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLD SVRKKVAVIQPPEYNRFALSFGHIF
	550 560 570 580 590 600
a128-1.pep	610 620 630 640 650 660
	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKA FRG
m128-1	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKA FRG
	610 620 630 640 650 660
a128-1.pep	670 679
	REPSIDALLRHSGFDNAAX
m128-1	REPSIDALLRHSGFDNAVX
	670

206

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3097>:  
m206.seq

m206 . pep . .

```

1 MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
51 QGSQELMLHS LGLIGTPYKQ GGSSTATGPD CSGMIQFVYK NALNVKLPRT
101 ARDMAAASRK IPDSRXKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTEFE*

```

g206.seq

1	atgtttttccc	ccgacaaaac	ccttttctct	tgtctcggcg	cactgctcct
51	cgcctcatgc	ggcacgacct	ccggcaaaca	ccgccaaccg	aaacccaaaac
101	agacagtcgc	gcaaatccaa	gccgtccgca	tcagccacat	cggcgcgaca
151	caaggctcgc	aggaactcat	gctccacagc	ctcggactca	tcggcacgcc
201	ctacaaatgg	ggcggcgagca	gcaccgcaac	cggcttcgac	tgcagcggga
251	tgattcaatt	ggtttacaaa	aacgcctca	acgtcaagct	gccgcgcacc
301	gcccgcgaca	tggcggcggc	aagccgcaaa	atccccgaca	gccgcctcaa
351	ggcgcggcag	atcgtattct	tcaacaccgg	cggcgcacac	cgctactcac
401	acgtcggact	ctacatctgc	aacggcgaat	tcattcatgc	ccccggcagc
451	ggcaaaaacca	tcaaaaccga	aaaactctcc	acaccgtttt	acgccaaaaa
501	ctaccttgga	gcgcatacgt	tttttacaqa	atqa	

g206 . pep

1 MFSPDKTLFL CLGALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIGRT  
51 QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQLVYK NALNVKLPRT  
101 ARDMAAASRK IPDSRLKAGD IVFFNTGGAH RYSHVGLYIG NGEFIHAPGS  
151 GKTIKTEKLS TPFYAKNYLG AHTEFTE\*

m206/q206

BNSDOCID: <WO\_8957280A2 | >

130 140 150 160 170

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3101>:

```
a206.seq
1  ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
51  CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
101 AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
151 CAAGGCTCGC AGGAACTCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
201 CTACAAATGG GGCGGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGGCA
251 TGATTCAATT CGTTTACAAA AACGCCCTCA ACGTCAAGCT GCCGCGCACC
301 GCCC GCGACA TGGCGGCGGC AAGCCGCAA ATCCCCGACA GCCGCCTTAA
351 GGCCGGGCGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGCTACTCAC
401 ACGTCGGACT CTATATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
451 GGCAAACCA TCAAAACCGA AAAACTCTCC ACACCGTTTT ACGCCAAAAA
501 CTACCTCGGC GCACATACTT TCTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3102; ORF 206.a>:

```
a206.pep
1  MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPR
101 ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

m206/a206 ORFs 206 and 206.a showed a 99.4% identity in 177 aa overlap

	10	20	30	40	50	60
m206.pep	MFPPDKTLFLCLLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
a206	MFPPDKTLFLCLLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m206.pep	LGLIGTPYKWGGSSTATGFD CSGMIQFVYK NALNVKLPR TARDMAAASRKIPDSRXKAGD					
a206	LGLIGTPYKWGGSSTATGFD CSGMIQFVYK NALNVKLPR TARDMAAASRKIPDSRLKAGD					
	70	80	90	100	110	120
	130	140	150	160	170	
m206.pep	LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX					
a206	LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX					
	130	140	150	160	170	

287

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3103>:

```
m287.seq
1  ATGTTTAAAC GCAGCGTAAT CGCAATGGCT TGTATTTTGT CCCTTTCAGC
51  CTGCGGGGGC GGCGGTGGCG GATCGCCCGA TGTCAGTTCG GCGGACACGC
101 TGTCAAACCC TGCCGCCCCCT GTTGTCTCTG AAAAAGAGAC AGAGGCAAAG
151 GAAGATGCGC CACAGGCAGG TTCTCAAGGA CAGGGCGCGC CATCCGCACA
201 AGGCAGTCAA GATATGGCGG CGGTTTCGGA AGAAAATACA GGCAATGGCG
251 GTGCGGTAAC AGCGGATAAT CCCAAAATG AAGACGAGGT GGCACAAAAT
301 GATATGCCGC AAAATGCCGC CGGTACAGAT AGTTCGACAC CGAATCACAC
351 CCCGGATCCG AATATGCTTG CCGGAAATAT GGAAAATCAA GCAACGGATG
401 CCGGGGAATC GTCTCAGCCG GCAAACCAAC CGGATATGGC AAATGCGGGC
451 GACGGAATGC AGGGGGACGA TCCGTCGGCA GGCGGGCAAA ATGCCGGCAA
501 TACGGCTGCC CAAGGTGCAA ATCAAGCCGG AAACAATCAA GCCGCCGGTT
551 CTTGAGATCC CATCCCCGCG TCAAACCTG CACCTGCGAA TGGCGGTAGC
```

```

601 AATTTTGGAA GGGTTGATTT GGCTAATGGC GTTTTGATTG ACGGGCCGTC
651 GCAAAATATA ACGTTGACCC ACTGTAAAGG CGATTCTTGT AGTGGCAATA
701 ATTTCTTGGG TGAAGAAGTA CAGCTAAAAT CAGAATTGGA AAAATTAAAT
751 GATGCAGACA AAATAAGTAA TTACAAGAAA GATGGGAAGA ATGATAAATT
801 TGTCGGTTTG GTTGCCGATA GTGTGCAGAT GAAGGGAATC AATCAATATA
851 TTATCTTTTA TAAACCTAAA CCCACTTCAT TTGCGCGATT TAGGCGTTC
901 GCACGGTCGA GCGGGTCGCT TCCGGCCGAG ATGCCGCTGA TTCCCGTCAA
951 TCAGGCGGAT ACGCTGATTG TCGATGGGGA AGCGGTCAGC CTGACGGGGC
1001 ATTCGGGCAA TATCTTCGCG CCCGAAGGGA ATTACCGGTA TCTGACTTAC
1051 GGGGCGGAAA AATTGCCCGG CGGATCGTAT GCCCTTCGTG TTCAAGGCGA
1101 ACCGGCAAAA GGCGAAATGC TTGCGGGCGC GGCCGTGTAC AACGGCGAAG
1151 TACTGCATTT CCATACGGAA AACGGCCGTC CGTACCCGAC CAGGGGCAAG
1201 TTTGCCGCAA AAGTCGATTT CGGCAGCAA TCTGTGGACG GCATTATCGA
1251 CAGCGGCGAT GATTTGCATA TGGGTACGCA AAAATTCAAA GCCGCCATCG
1301 ATGGAAACGG CTTTAAGGGG ACTTGGACGG AAAATGGCAG CGGGGATGTT
1351 TCCGGAAAGT TTTACGGCCC GGCCGGCGAG GAAGTGGCGG GAAAATACAG
1401 CTATCGCCCG ACAGATGCGG AAAAGGGCGG ATTCGGCGTG TTTGCCGCGA
1451 AAAAAGAGCA GGATTGA

```

This corresponds to the amino acid sequence <SEQ ID 3104; ORF 287>:

m287.pep

```

1 MFKRSVIAMA CIFALSACGG GGGGSPDVKS ADTLSKPAAP VVSEKETEAK
51 EDAPQAGSQ QGAPSAQGSQ DMAAVSEENT GNGGAVTADN PKNEDEVAQN
101 DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP ANQPDMANAA
151 DGMQGDPSA GGQNAAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
201 NFGRVDLANG VLIDGPSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS
251 DADKISNYKK DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRS
301 ARSRRSLPAE MPLIPVNQAD TLIVDGEAVS LTGHSNIFA PEGNYRYLTY
351 GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY NGEVLHFHTE NGRPYPTRGR
401 FAAKVDGSK SVDGIIDSGD DLHMGTKQFK AAIDGNGFKG TWTENGSGDV
451 SGKFGYPAGE EVAGKYSYRP TDAEKGFGV FAGKEQD*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3105>:

g287.seq

```

1 atgtttaaac gcagtgtgat tgcaatggct tgtatttttc ccccttcagc
51 ctgtgggggc ggcggtggcg gatcgcccg tgtcaagtcg gcggaacgcg
101 ctgcaaaacc ggccgcccc gttgttgctg aaaaatgcgg ggaaggggtg
151 ctgccgaaag aaaagaaaga tgaggaggca gcggcggtg cgcgcgaagc
201 cgatacgag gacgcaaccg ccggagaagg cagccaagat atggcggcag
251 tttcggcaga aaatacaggc aatggcggtg cggcaacaac ggacaacccc
301 aaaaatgaag acgcgggggc gcaaaatgat atgccgcaa atgccgccga
351 attccgcaaat caaacaggga acaaccaacc cgccggttct tcagattccg
401 ccccgcgctc aaaccctgcc cctgcgaatg gcggtagcga ttttggaagg
451 acgaacgtgg gcaattctgt tgtgattgac ggaccgtcgc aaaatataac
501 gttgaccac tgtaaaggcg attcctgtaa tggtgataat ttattggatg
551 aagaagcacc gtcaaaatca gaatttgaaa aattaagtga tgaagaaaaa
601 attaaagcat ataaaaaaga cgagcaacgg gagaattttg tcggtttggt
651 tgctgacagg gtaaaaaagg atggaactaa caaatatata atcttctata
701 cggacaaaacc acctactcgt tctgcacggt cgaggaggct gcttccggcc
751 gagattccgc tgattcccg caatcaggcc gatacgctga ttgtggatgg
801 ggaagcggtc agcctgacgg ggcattccgg caatatcttc gcgccgaag
851 ggaattaccg gtatctgact tacggggcgg aaaaattgcc cggcggtcgc
901 tatgccctcc gtgtgcaagg cgaaccggca aaaggcgaaa tgcttgttgg
951 cacggccgtg tacaacggcg aagtgtgca tttccatatg gaaaacggcc
1001 gtccgtaccc gtccggaggc aggtttgccg caaaagtcga tttcggcagc
1051 aaatctgtgg acggcattat cgacagcggc gatgatattgc atatgggtac
1101 gcaaaaaattc aaagccgcca tcgatggaaa cggctttaag gggacttgga
1151 cggaaaatgg cggcggggat gtttccggaa ggttttacgc cccggccggc
1201 gaggaagtgg cgggaaaata cagctatcgc ccgacagatg ctgaaaaggc
1251 cggattcggc gtgtttgccg gcaaaaaaga tcgggattga

```

This corresponds to the amino acid sequence <SEQ ID 3106; ORF 287.ng>:

g287.pep

```

1 MFKRSVIAMA CIFPLSACGG GGGGSPDVKS ADTPSKPAAP VVAENAGEGV

```

```

51 LPKEKKDEEA AGGAPQADTQ DATAGEGSQD MAAVSAENTG NGGAATTDNP
101 KNEDAGAQN DMPQNAAESAN QTGNNQPAGS SDSAPASNPA PANGGSDFGFR
151 TNVGNVSVVID GPSQNTLTH CKGDSCNGDN LLDEEAPSKS EFEKLSDEEK
201 IKRYKKDEQR ENFVGLVADR VKKDGTNKYI IFYTDKPPTR SARSRRLPA
251 EIPLIPVNQA DTLIVDGEAV SLTGHSGNIF APEGNYRYLT YGAEKLPGGS
301 YALRVQGEPA KGEMLVGTAV YNGEVLHFHM ENGRPYPSGG RFAAKVDFGS
351 KSV DGIIDSG DDLHMGTKQF KAAIDGNGFK GTWTENGGGD VSGRFYGPAG
401 EEVAGKYSYR PTDAEKGGFG VFAGKKDRD*

```

**m287/g287** ORFs 287 and 287.ng showed a 70.1% identity in 499 aa overlap

```

          10      20      30      40
m287.pep  MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSE-----KETE
          10      20      30      40      50      60
g287      MFKRSVIAMACIFPLSACGGGGGGSPDVKSADTPSKPAAPVVAENAGEGVLPKEKKDEEA

          50      60      70      80      90      100      109
m287.pep  KEDAPQAGSQGQAGPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQN DMPQNAAGT
          10      20      30      40      50      60
g287      AGGAPQADTQD--ATAGEGSQDMAAVSAENTGNGGAATTDNPKNEDEVAQN DMPQNAAGT

          110      120      130      140      150      160      169
m287.pep  DSSTPNHTPDNMLAGNMENQATDAGESSQFANQPDMANAADGMQGGDDPSAGGQNAAGNTA
          110      120      130      140      150      160      169
g287      -----

          170      180      190      200      210      220      229
m287.pep  AQGANQAGNNQAAGSSDPIPASNPAPANGGNSFGRVDLANGVLIDGPSQNTLTHCKGDS
          120      130      140      150      160      170
g287      -ESANQTGNNQPAGSSDSAPASNPAANGGSDFGRTNVGNVSVVIDGPSQNTLTHCKGDS

          230      240      250      260      270      280      289
m287.pep  CSGNNFLDEEVQLKSEFEKLSADKISNYKKDKGNDKFVGLVADSVQMKGINQYIIFYKP
          180      190      200      210      220      230
g287      CNGDNLLEDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVKKDGTNKYIIFYTD

          290      300      310      320      330      340      349
m287.pep  KPTSFAFRFRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLT
          240      250      260      270      280      290
g287      KPPT-----RSARSRRSLPAEIPVQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLT

          350      360      370      380      390      400      409
m287.pep  YGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGS
          300      310      320      330      340      350
g287      YGAEKLPGGSYALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGRFAAKVDFGS

          410      420      430      440      450      460      469
m287.pep  KSV DGIIDSGDDLHMGTKQFKAAIDGNGFKGTWTENSGDVSFGFYGPAGEEVAGKYSYR
          360      370      380      390      400      410
g287      KSV DGIIDSGDDLHMGTKQFKAAIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYR

          470      480      489
m287.pep  PTDAEKGGFGVFAGKKEQDX

```



g287  
 |||||:|  
 PTDAEKGFGVFAGKKDRDX  
 420 430

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3107>:

a287.seq  
 1 ATGTTTAAAC GCAGTGTGAT TGCAATGGCT TGTATTGTTG CCCTTTCAGC  
 51 CTGTGGGGGC GCGGTGGCG GATCGCCCGA TGTAAAGTCG GCGGACACGC  
 101 TGTCAAACC TGCCGCCCTT GTTGTACTG AAGATGTCGG GGAAGAGGTG  
 151 CTGCCGAAAG AAAAGAAAGA TGAGGAGGCG GTGAGTGGTG CGCCGCAAGC  
 201 CGATACGCAG GACGCAACCG CCGGAAAAGG CCGTCAAGAT ATGGCGGCAG  
 251 TTTCCGGCAGA AAATACAGGC AATGGCGGTG CCGCAACAAC GGATAATCCC  
 301 GAAAATAAAG ACGAGGGACC GCAAAATGAT ATGCCGCAAA ATGCCCGCGA  
 351 TACAGATAGT TCGACACCGA ATCACACCCC TGCACCGAAT ATGCCAACCA  
 401 GAGATATGGG AAACCAAGCA CCGGATGCCG GGAATCGGC ACAACCGGCA  
 451 AACCAACCGG ATATGGCAAA TCGCGCGGAC GGAATGCAGG GGGACGATCC  
 501 GTCGGCAGGG GAAAATGCCG GCAATACGGC AGATCAAGCT GCAAATCAAG  
 551 CTGAAAACAA TCAAGTCGGC GGCTCTCAA ATCCTGCCTC TTCAACCAAT  
 601 CCTAACGCCA CGAATGGCGG CAGCGATTTT GGAAGGATAA ATGTAGCTAA  
 651 TGGCATCAAG CTTGACAGCG GTTCGGAAAA TGTAACGTTG ACACATTGTA  
 701 AAGACAAAGT ATGCGATAGA GATTCTTAG ATGAAGAAGC ACCACCAAAA  
 751 TCAGAATTG AAAAATTAAG TGATGAAGAA AAAATTAATA AATATAAAAA  
 801 AGACGAGCAA CGAGAGAATT TTGTGCGTTT GGTGCTGAC AGGGTAGAAA  
 851 AGAATGGAAC TAACAAATAT GTCATCATT ATAAAGACAA GTCCGCTTCA  
 901 TCTTCATCTG CGCGATTCAG GCGTCTGCA CCGTCGAGGC GGTCGCTTCC  
 951 GGCCGAGATG CCGCTGATTC CCGTCAATCA GCGGATACG CTGATTGTCC  
 1001 ATGGGGAAGC GGTCAGCCTG ACGGGGCATT CCGGCAATAT CTTCCGCGCC  
 1051 GAAGGGAATT ACCGGTATCT GACTTACGGG GCGGAAAAAT TGTCCGCGCG  
 1101 ATCGTATGCC CTCAGTGTGC AAGGCGAACC GGCAAAAGGC GAAATGCTTG  
 1151 CGGGCACGGC CGTGTACAAC GGCGAAGTGC TGCAATTCCA TATGGAACAA  
 1201 GGCCGTCCGT CCCCCTCCGG AGGCAGGTTT GCCGCAAAAG TCGATTTCGG  
 1251 CAGCAAATCT GTGGACGGCA TTATCGACAG CCGCGATGAT TTGCATATGG  
 1301 GTACGCAAAA ATTCAAAGCC GTTATCGATG GAAACGGCTT TAAGGGGACT  
 1351 TGGACGGAAA ATGGCGGCGG GGATGTTTCC GGAAGGTTTT ACGGCCCCGG  
 1401 CGGCGAAGAA GTGGCGGGAA AATACAGCTA TCGCCCGACA GATGCGGAAA  
 1451 AGGGCGGATT CGGCGTGTTC GCCGGCAAAA AAGAGCAGGA TTGA

This corresponds to the amino acid sequence <SEQ ID 3108; ORF 287.a>:

a287.pep  
 1 MFKRSVIAMA CIVALSAACG GGGGSPDVKS ADTLSKPAAP VVTEDVGEEV  
 51 LPKEKKDEEA VSGAPQADTQ DATAGKGGQD MAAVSAENTG NGGAATTDNPF  
 101 ENKDEGPQND MPQNAADTDS STPNHTPAPN MPTRDMGNQA PDAGESAQPA  
 151 NQPDMAAAD GMQDDPSAG ENAGNTADQA ANQAENNQVG GSQNPASSTN  
 201 PNATNGGSDF GRINVANGIK LDGSENVTL THCKDKVCDR DFLDEEAPPK  
 251 SEFEKLSDEE KINKYKKDEQ RENFVGLVAD RVEKNGTNKY VIYKDKSAS  
 301 SSSARFRSA RSRRSLPAEM PLIPVNQADT LIVDGEAVSL TGHSGNIFAP  
 351 EGNRYRLTYG AEKLSGGSYA LSVQGEPAKG EMLAGTAVYN GEVLHFHMFN  
 401 GRPSPSGGRF AAKVDFGSKS VDGIIIDSGD LHMGTQKFKF VIDGNNGFKGT  
 451 WTENGGGDVS GRFYGPAGEE VAGKYSYRPT DAEKGGFGVF AGKKEQD\*

m287/a287 ORFs 287 and 287.a showed a 77.2% identity in 501 aa overlap

	10	20	30	40	49
m287.pep	MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSE-----KETEA				
a287	MFKRSVIAMACIVALSAACGGGGGGSPDVKSADTLSKPAAPVVTEDEVGEEVLPKEKKDEEA				
	10	20	30	40	50
	50	60	70	80	90
m287.pep	KEDAPQAGSQGQGPSAQGSQDMAVSEENTGNGGAVTADNPKNEDEVAQNMPQNAAGT				
a287	VSGAPQADTQ--DATAGKGGQDMAVSAENTGNGGAATTDNPNKDEGPQNDMPQNAADT				
	70	80	90	100	110

406

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3109>:

m406.seq

1	ATGCAAGCAC	GGCTGCTGAT	ACCTATTCTT	TTTTCAGTTT	TTATTTTATC
51	CGCCTGCGGG	ACACTGACAG	GTATTCCATC	GCATGGCGGA	GGTAAACGCT
101	TTGCGGTCGA	ACAAGAACTT	GTGGCCGCTT	CTGCCAGAGC	TGCCGTTAAA
151	GACATGGATT	TACAGGCATT	ACACGGACGA	AAAGTTGCAT	TGTACATTGC
201	CACATGGGGC	GACCAAGGTT	CAGGCAGTTT	GACAGGGGGT	CGCTACTCCA
251	TTGATGCACT	GATTTCGTGC	GAATACATAA	ACAGCCCTGC	CGTCCGTACC
301	GATTACACCT	ATCCACGTTA	CGAAACCAAC	GCTGAAACAA	CATCAGGCGT
351	TTTGACAGGT	TTAACCACTT	CTTTATCTAC	ACTTAATGCC	CCTGCACCTT
401	CTCGCACCCA	ATCAGACCGT	AGCGGAAGTA	AAAGCAGTCT	GGGCTTAAAT
451	ATTGGCGGGA	TGGGGGATTA	TCGAAATGAA	ACCTTGACGA	CTAACCCGCG
501	CGACACTGCC	TTTCTTTCCC	ACTTGGTATA	GACCGTATTT	TTCTGCGCG
551	GCATAGACGT	TGTTTCTCCT	GCCAAATGCC	ATACAGATGT	GTTTATTAAC
601	ATCGACGTAT	TCGGAACGAT	ACGCAACAGA	ACCGAAATGC	ACCTATACAA
651	TGCCGAAACA	CTGAAAGCCC	AAACAAAAC	GGAATATTTT	GCAGTAGACA

```

701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACGGAAG GATTAATGGT CGATTTCTCC GATATCCGAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GTAGTGCGAC AACATAGACA
951 AGGACAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3110; ORF 406>:

m406.pep

```

1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSGLTG LTTSLSLTLNA PALSRTQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRNTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSIGIK PTEGLMVDFS DIRPYGNHTG NSAPSVEADN
301 SHEGYGYSDE VVRQHRQGP *

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3111>:

g406.seq

```

1  ATGCGGGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG AACTGACAG GTATTCCATC GCATGGCGGA GGCAAACGCT
101 TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251 TTGATGCACT GATTCGCGGC GAATACATAA ACAGCCCTGC CGTCCGCACC
301 GATTACACCT ATCCGCGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACGGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA GGAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CCAACCCGCG
501 CGACACTGCC TTCTTTCCC ACTTGGTGCA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACAGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAACT GGAATATTT GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATCAAA CCGACGGAAG GATTGATGGT CGATTTCTCC GATATCCAAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCGAC AACATAGACA
951 AGGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3112; ORF 406>:

g406.pep

```

1  MRARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSGLTG LTTSLSLTLNA PALSRTQSDG SGSRSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRNTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSIGIK PTEGLMVDFS DIQPYGNHTG NSAPSVEADN
301 SHEGYGYSDE AVRQHRQGP *

```

ORF 406 shows 98.8% identity over a 320 aa overlap with a predicted ORF (ORF406.a) from *N. gonorrhoeae*:

g406/m406

```

              10      20      30      40      50      60
g406.pep  MRARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAAVKDMDLQALHGR
           |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
m406      MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAAVKDMDLQALHGR

```

	10	20	30	40	50	60
g406.pep	70	80	90	100	110	120
	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
m406	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
	70	80	90	100	110	120
g406.pep	130	140	150	160	170	180
	LTTSLSTLNAPALSRTQSDGSGSRSSLGLNIGGMGDYRNETLTNPRDTAFLSHLVQTVF					
m406	LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTNPRDTAFLSHLVQTVF					
	130	140	150	160	170	180
g406.pep	190	200	210	220	230	240
	FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
m406	FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
	190	200	210	220	230	240
g406.pep	250	260	270	280	290	300
	IKPKTNAFEAAAYKENYALWMGPYKVS KG IKPTEGLMVDFSDIQPYGNHTGNSAPSVEADN					
m406	IKPKTNAFEAAAYKENYALWMGPYKVS KG IKPTEGLMVDFSDIRPYGNHTGNSAPSVEADN					
	250	260	270	280	290	300
g406.pep	310	320				
	SHEGYGYSDEAVRQHRQGQPX					
m406	SHEGYGYSDEVVRQHRQGQPX					
	310	320				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3113>:

```

a406.seq
1  ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG ACACTGACAG GTATTCCATC GCATGGCGGA GGTAACGCT
101 TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGACATTGC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGT CGTACTCCA
251 TTGATGCACT GATTTCGTGGC GAATACATAA ACAGCCCTGC CGTCCGTACC
301 GATTACACCT ATCCACGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACGGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAAC GGAATATTT GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGACCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACAGAAG GATTAATGGT CGATTTCTCC GATATCCAAC
851 CATACGGCAA TCATATGGGT AACTCTGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCAC GACATAGACA
951 AGGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3114; ORF 406.a>:

```

a406.pep
1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSGSLTG RYSIDALIRG EYINSPAVRT
101 DYTPRYETT AETTSGGLTG LTTSLSTLNA PALSRQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA

```

251 AYKENYALWM GPYKVSCKGIK PTEGLMVDFS DIQPYGNHMG NSAPSVEADN  
 301 SHEGYGYSDE AVRRHRQGQP \*

m406/a406 ORFs 406 and 406.a showed a 98.8% identity in 320 aa overlap

	10	20	30	40	50	60
m406.pep	MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA AVKMDLQALHGR					
a406	MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA AVKMDLQALHGR					
	10	20	30	40	50	60
	70	80	90	100	110	120
m406.pep	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETSSGGLTG					
a406	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETSSGGLTG					
	70	80	90	100	110	120
	130	140	150	160	170	180
m406.pep	LTTSLSTLNAPALSRTQSDGSGSKSSGLNIGMGDYRNETLTTNPRDTAFLSHLVQTVF					
a406	LTTSLSTLNAPALSRTQSDGSGSKSSGLNIGMGDYRNETLTTNPRDTAFLSHLVQTVF					
	130	140	150	160	170	180
	190	200	210	220	230	240
m406.pep	FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
a406	FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m406.pep	IKPKTNAFEAAAYKENYALWMGPYKVSCKGIKPTGLMVDFS DIRPYGNHTGNSAPSVEADN					
a406	IKPKTNAFEAAAYKENYALWMGPYKVSCKGIKPTGLMVDFS DIQPYGNHMGNSAPSVEADN					
	250	260	270	280	290	300
	310	320				
m406.pep	SHEGYGYSDEVVRQHRQGQPX					
a406	SHEGYGYSDEAVRRHRQGQPX					
	310	320				

## EXAMPLE 2

### Expression of ORF 919

The primer described in Table 1 for ORF 919 was used to locate and clone ORF 919. The predicted gene *919* was cloned in pET vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 919-His fusion protein purification. Mice were immunized with the purified 919-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; PP, purified protein, TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm

that 919 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 919 are provided in Figure 10. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 919 and the amino acid sequence encoded thereby is provided in Example 1.

### EXAMPLE 3

#### Expression of ORF 279

The primer described in Table 1 for ORF 279 was used to locate and clone ORF 279. The predicted gene 279 was cloned in pGex vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 279-GST purification. Mice were immunized with the purified 279-GST and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 279 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 279 are provided in Figure 11. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 279 and the amino acid sequence encoded thereby is provided in Example 1.

### EXAMPLE 4

#### Expression of ORF 576 and 576-1

The primer described in Table 1 for ORF 576 was used to locate and clone ORF 576. The predicted gene 576 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 576-GST fusion protein purification. Mice were immunized with the purified 576-GST and sera

were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B).. These experiments confirm that ORF 576 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 576 are provided in Figure 12. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 576 and the amino acid sequence encoded thereby is provided in Example 1.

#### EXAMPLE 5

##### Expression of ORF 519 and 519-1

The primer described in Table 1 for ORF 519 was used to locate and clone ORF 519. The predicted gene 519 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 519-His fusion protein purification. Mice were immunized with the purified 519-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 519 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 13. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 519 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 6

Expression of ORF 121 and 121-1

The primer described in Table 1 for ORF 121 was used to locate and clone ORF 121. The predicted gene *121* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 121-His fusion protein purification. Mice were immunized with the purified 121-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 121 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 121 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 121 are provided in Figure 14. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 121 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 7

Expression of ORF 128 and 128-1

The primer described in Table 1 for ORF 128 was used to locate and clone ORF 128. The predicted gene *128* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 128-His purification. Mice were immunized with the purified 128-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D) and ELISA assay (panel E). Results show that 128 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 128 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 128 are provided in Figure 15. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J.*



*Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 128 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 8

### Expression of ORF 206

The primer described in Table 1 for ORF 206 was used to locate and clone ORF 206. The predicted gene 206 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 206-His purification. Mice were immunized with the purified 206-His and sera were used for Western blot analysis (panel B). It is worth noting that the immunoreactive band in protein extracts from meningococcus is 38 kDa instead of 17 kDa (panel A). To gain information on the nature of this antibody staining we expressed ORF 206 in *E. coli* without the His-tag and including the predicted leader peptide. Western blot analysis on total protein extracts from *E. coli* expressing this native form of the 206 protein showed a reactive band at a position of 38 kDa, as observed in meningococcus. We conclude that the 38 kDa band in panel B) is specific and that anti-206 antibodies, likely recognize a multimeric protein complex. In panel C is shown the FACS analysis, in panel D the bactericidal assay, and in panel E) the ELISA assay. Results show that 206 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 206 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 519 are provided in Figure 16. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 206 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 9

Expression of ORF 287

The primer described in Table 1 for ORF 287 was used to locate and clone ORF 287. The predicted gene 287 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 287-GST fusion protein purification. Mice were immunized with the purified 287-GST and sera were used for FACS analysis (panel B), bactericidal assay (panel C), and ELISA assay (panel D). Results show that 287 is a surface-exposed protein. Symbols: M1, molecular weight marker. Arrow indicates the position of the main recombinant protein product (A). These experiments confirm that 287 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 287 are provided in Figure 17. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 287 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 10

Expression of ORF 406

The primer described in Table 1 for ORF 406 was used to locate and clone ORF 406. The predicted gene 406 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 406-His fusion protein purification. Mice were immunized with the purified 406-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 406 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 406 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 406 are provided in Figure 18. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al.

1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 406 and the amino acid sequence encoded thereby is provided in Example 1.

### EXAMPLE 11

Table 2 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 225 among different strains.

**Table 2**

#### **225 gene variability: List of used *Neisseria* strains**

<b>Identification Strains number</b>	<b>Source / reference</b>
<b>Group B</b>	
zo01_225 NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo02_225 BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zo03_225 NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo04_225 297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zo05_225 1000	R. Moxon / Seiler <i>et al.</i> , 1996
zo06_225 BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zo07_225 BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zo08_225 528	R. Moxon / Seiler <i>et al.</i> , 1996
zo09_225 NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zo10_225 BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zo11_225 NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zo12_225 NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zo13_225 NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zo14_225 NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zo15_225 SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zo16_225 NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zo17_225 NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zo18_225 BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zo19_225 BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zo20_225 44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zo21_225 MC58	R. Moxon
zo96_225 2996	Our collection
<b>Group A</b>	
zo22_225 205900	R. Moxon
zo23_225 F6124	R. Moxon
z2491 Z2491	R. Moxon / Maiden <i>et al.</i> , 1998
<b>Group C</b>	
zo24_225 90/18311	R. Moxon
zo25_225 93/4286	R. Moxon

**Others**

zo26\_225 A22 (group W) R. Moxon / Maiden *et al.*, 1998  
 zo27\_225 E26 (group X) R. Moxon / Maiden *et al.*, 1998  
 zo28\_225 860800 (group Y) R. Moxon / Maiden *et al.*, 1998  
 zo29\_225 E32 (group Z) R. Moxon / Maiden *et al.*, 1998

**Gonococcus**

zo32\_225 Ng F62 R. Moxon / Maiden *et al.*, 1998  
 zo33\_225 Ng SN4 R. Moxon  
  
 fa1090 FA1090 R. Moxon

**References:**

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
 Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

>FA1090 <SEQ ID 3115>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS  
 GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
 RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKNDPSRFLN\*

Z2491 <SEQ ID 3116>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRVPARRAGNA  
 DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF  
 MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF  
 IHAHPTGKNIEITSLSHKYWSGKYAFARRVKNDPSRFLN\*

Z001\_225 <SEQ ID 3117>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
 DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAHPTGKNIEITSLSHKYWSGKYAFARR  
 VKNDPSRFLN\*

Z002\_225 <SEQ ID 3118>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
 DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAHPTGKNIEITSLSHKYWSGKYAFARR  
 VKNDPSRFLN\*

Z003\_225 <SEQ ID 3119>  
 MDSFFKPAVWAVLWLMFAVRLALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
 DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAHPTGKNIEITSLSHKYWSGKYAFARR  
 VKNDPSRFLN\*

Z004\_225 &lt;SEQ ID 3120&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z005\_225 &lt;SEQ ID 3121&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z006\_225 &lt;SEQ ID 3122&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z007\_225 &lt;SEQ ID 3123&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z008\_225 &lt;SEQ ID 3124&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z009\_225 &lt;SEQ ID 3125&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z010\_225 &lt;SEQ ID 3126&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z011\_225 &lt;SEQ ID 3127&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF  
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF  
IAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z012\_225 &lt;SEQ ID 3128&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z013\_225 &lt;SEQ ID 3129&gt;

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z014\_225 <SEQ ID 3130>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z015\_225 <SEQ ID 3131>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCS  
GFMQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
RFIAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z016\_225 <SEQ ID 3132>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z017\_225 <SEQ ID 3133>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z018\_225 <SEQ ID 3134>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z019\_225 <SEQ ID 3135>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z020\_225 <SEQ ID 3136>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGSAMGLNEQPVLVNRVPARRAGNA  
DELIGNAMGLNEQPVLVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF  
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF  
IAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z021\_225 <SEQ ID 3137>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z022\_225 <SEQ ID 3138>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG  
NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR

SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z023\_225 <SEQ ID 3139>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z024\_225 <SEQ ID 3140>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z025\_225 <SEQ ID 3141>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z026\_225 <SEQ ID 3142>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z027\_225 <SEQ ID 3143>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z028\_225 <SEQ ID 3144>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z029\_225 <SEQ ID 3145>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKN DPSRFLN\*

Z032\_225 <SEQ ID 3146>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS  
GFMQHIKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN\*

Z033\_225 <SEQ ID 3147>

MDSFFKPAVWAVLWLMFAVRSALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS  
GFMQHIKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRIKKN DPSRFLN\*

Z096\_225 <SEQ ID 3148>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG  
 NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA  
 DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHI FKRAMGINLPRTSAEQARMGTPVAR  
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
 VKKNDPSRFLN\*

Figure 19 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 225, further confirming its utility as an antigen for both vaccines and diagnostics.

### EXAMPLE 12

Table 3 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 235 among different strains.

**Table 3**

#### 235 gene variability: List of used *Neisseria* strains

Identification Strains number	Reference
<b>Group B</b>	
gnmzq01 NG6/88	Seiler <i>et al.</i> , 1996
gnmzq02 BZ198	Seiler <i>et al.</i> , 1996
gnmzq03 NG3/88	Seiler <i>et al.</i> , 1996
gnmzq04 1000	Seiler <i>et al.</i> , 1996
gnmzq05 1000	Seiler <i>et al.</i> , 1996
gnmzq07 BZ169	Seiler <i>et al.</i> , 1996
gnmzq08 528	Seiler <i>et al.</i> , 1996
gnmzq09 NGP165	Seiler <i>et al.</i> , 1996
gnmzq10 BZ133	Seiler <i>et al.</i> , 1996
gnmzq11 NGE31	Seiler <i>et al.</i> , 1996
gnmzq13 NGE28	Seiler <i>et al.</i> , 1996
gnmzq14 NGH38	Seiler <i>et al.</i> , 1996
gnmzq15 SWZ107	Seiler <i>et al.</i> , 1996
gnmzq16 NGH15	Seiler <i>et al.</i> , 1996
gnmzq17 NGH36	Seiler <i>et al.</i> , 1996
gnmzq18 BZ232	Seiler <i>et al.</i> , 1996
gnmzq19 BZ83	Seiler <i>et al.</i> , 1996
gnmzq21 MC58	Virji <i>et al.</i> , 1992
<b>Group A</b>	
gnmzq22 205900	Our collection



gnmzq23 F6124 Our collection  
z2491 Z2491 Maiden *et al.*, 1998

### Group C

gnmzq24 90/18311 Our collection  
gnmzq25 93/4286 Our collection

### Others

gnmzq26 A22 (group W) Maiden *et al.*, 1998  
gnmzq27 E26 (group X) Maiden *et al.*, 1998  
gnmzq28 860800 (group Y) Maiden *et al.*, 1998  
gnmzq29 E32 (group Z) Maiden *et al.*, 1998  
gnmzq31 *N. lactamica* Our collection

### Gonococcus

gnmzq32 Ng F62 Maiden *et al.*, 1998  
gnmzq33 Ng SN4 Our collection

fa1090 FA1090 Dempsey *et al.* 1991

### References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.  
Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279  
Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3149>

MKPLILGLAAVLALSACQVRKAPDLDTYSFKESKPASILVVPPLNESPDVNGTWGMLAST  
AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK\*

GNMZQ01 <SEQ ID 3150>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANNLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ02 <SEQ ID 3151>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ03 <SEQ ID 3152>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST

AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ04 <SEQ ID 3153>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ05 <SEQ ID 3154>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ07 <SEQ ID 3155>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ08 <SEQ ID 3156>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ09 <SEQ ID 3157>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
AEP LSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ10 <SEQ ID 3158>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ11 <SEQ ID 3159>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ13 <SEQ ID 3160>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ14 <SEQ ID 3161>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ15 <SEQ ID 3162>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ16 <SEQ ID 3163>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ17 <SEQ ID 3164>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ18 <SEQ ID 3165>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ19 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ21 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ22 <SEQ ID 3167>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ23 <SEQ ID 3168>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ24 <SEQ ID 3169>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ25 <SEQ ID 3170>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ26 <SEQ ID 3171>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ27 <SEQ ID 3172>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ28 <SEQ ID 3173>  
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ29 <SEQ ID 3174>  
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ31 <SEQ ID 3175>  
 MKPLILGLAAVLALSACQVQKAPDFDYTAFKESKPASILVVPPLNESPDVNGTWGMLAST  
 AEPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITITEYGTS  
 YQILDSVTTVSARARLVDSRNGKVLWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT  
 DRGYQVSKAAAYDLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ32 <SEQ ID 3176>  
 MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
 AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK\*

GNMZQ33 <SEQ ID 3177>  
 MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
 AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK\*

Z2491 <SEQ ID 3178>  
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

Figure 20 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 235, further confirming its utility as an antigen for both vaccines and diagnostics.

### EXAMPLE 13

Table 4 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 287 among different strains.

**Table 4**

#### **287 gene variability: List of used *Neisseria* strains**

<b>Identification Strains number</b>	<b>Reference</b>
--	------------------

**Group B**

287_2	BZ198	Seiler <i>et al.</i> , 1996
287_9	NGP165	Seiler <i>et al.</i> , 1996
287_14	NGH38	Seiler <i>et al.</i> , 1996
287_21	MC58	Virji <i>et al.</i> , 1992

**Group A**

z2491	Z2491	Maiden <i>et al.</i> , 1998
-------	-------	-----------------------------

**Gonococcus**

fa1090	FA1090	Dempsey <i>et al.</i> 1991
--------	--------	----------------------------

**References:**

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
 Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.  
 Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279  
 Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

287\_14 <SEQ ID 3179>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG  
 QGAPSAQGGQDMAAVSEENTGNNGGAAATDKPKNEDEGAQNMPQNAADTDSLTPNHTPAS  
 NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNO  
 TAGSQNPASSTNPSATNSGGDFGRNTVGNNSVVIDGPSQNTLTHCKGDCSGNNFLDEEV  
 QLKSEFEKLSADAKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIIFYKPKPTSFAFR  
 FRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLP  
 GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSRGRFAAKVDFGSKSVDGII  
 DSGDGLHMGTKQFKAAIDGNGFKGTWTENGGDVSFGFYGPAGEEVAGKYSYRPTDAEKG  
 GFGVFAGKKEQD\*

287\_2 <SEQ ID 3180>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG  
 QGAPSAQGGQDMAAVSEENTGNNGGAAATDKPKNEDEGAQNMPQNAADTDSLTPNHTPAS  
 NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNO  
 TAGSQNPASSTNPSATNSGGDFGRNTVGNNSVVIDGPSQNTLTHCKGDCSGNNFLDEEV  
 QLKSEFEKLSADAKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIIFYKPKPTSFAFR  
 FRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLP  
 GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSRGRFAAKVDFGSKSVDGII  
 DSGDGLHMGTKQFKAAIDGNGFKGTWTENGGDVSFGFYGPAGEEVAGKYSYRPTDAEKG  
 GFGVFAGKKEQD\*

287\_21. <SEQ ID 3181>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG  
 QGAPSAQGSQDMAAVSEENTGNNGGAVTADNPKNEDEVAQNMPQNAAGTDSSTPNHTPDP  
 NMLAGNMENQATDAGESSQPANQPDMAANAADGMQGGDDPSAGGQAGNTAAQGANQAGNNQ  
 AAGSSDPIPASNPAPANGGSNFRVVDLANGVLIDGPSQNTLTHCKGDCSGNNFLDEEV  
 QLKSEFEKLSADAKISNYKKDGKNDKFVGLVADSVQMKGINQYIIIFYKPKPTSFAFRFRS

ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSY  
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVDDGIIDSGD  
DLHMGTOQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGFGV  
FAGKKEQD\*

287\_9 <SEQ ID 3182>

MFKRSVIAMACIVALSAACGGGGGGSPDVKSADTLSPKPAAPVVTEDEVGEEVLPKEKKDEEA  
VSGAPQADTQDATAGKGGQDMAAVSAENTGNGGAATTDNPNKDEGPQNDMPQNAADTDS  
STPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMAAADGMQGGDDPSAGENAGNTADQA  
ANQAENNQVGGSONPASSTNPATNGGSDFGFRINVANGIKLDSGSENVTLTHCKDKVCDR  
DFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVEKNGTNKYVYIYKDKSAS  
SSSARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYG  
AEKLSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHFMENGRPSPSGGRFAAKVDFGSKS  
VDGIIDSGDDLHMGTOQKFKAAIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPT  
DAEKGFGVFAGKKEQD\*

FA1090 <SEQ ID 3183>

MFKRSVIAMACIFPLSACGGGGGGSPDVKSADTPSKPAPVVAENAGEGVLPKEKKDEEA  
AGGAPQADTQDATAGEGSDMAAVSAENTGNGGAATTDNPKNEDAGAQNMPQNAAESAN  
QTGNQNPAGSSDSAPASNPAPANGGSDFGRTNVGNSVVIDGPSQNTLTHCKGDSNGDN  
LLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVEKNGTNKYVYIYKDKSAS  
SARSRRSLPAEIPVQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGS  
YALRVQGEPAKGEMLVGTAVYNGEVLHFHFMENGRPSPSGGRFAAKVDFGSKSVDDGIIDSG  
DDLHMGTOQKFKAAIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPTDAEKGFGV  
VFAGKKDRD\*

Z2491 <SEQ ID 3184>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG  
QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNMPQNAAGTDSSTPNHTPDP  
NMLAGNMENQATDAGESSQPANQPDMAAADGMQGGDDPSAGGQNAAGNTAAQGANQAGNNQ  
AAGSSDPIPASNPAPANGGNSNFGRVLDLNGVLIDGPSQNTLTHCKGDSGSGNNFLDEEV  
QLKSEFEKLSADAKISNYKKDKGKDFVGLVADSVQMKGINQYIIFYKPKPTSFAFRFRS  
ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSY  
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVDDGIIDSGD  
DLHMGTOQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGFGV  
FAGKKEQD\*

Figure 21 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 287, further confirming its utility as an antigen for both vaccines and diagnostics.

#### EXAMPLE 14

Table 5 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 519 among different strains.

**Table 5**

#### **519 gene variability: List of used *Neisseria* strains**

Identification Strains	Source / reference
------------------------	--------------------

**number****Group B**

zv01_519	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv02_519	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zv03_519ass	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv04_519	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zv05_519	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zv06_519ass	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zv07_519	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zv11_519	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zv12_519	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zv18_519	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zv19_519	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zv20_519ass	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zv21_519ass	MC58	R. Moxon
zv96_519	2996	Our collection

**Group A**

zv22_519ass	205900	R. Moxon
z2491_519	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

**Others**

zv26_519	A22	(group W) R. Moxon / Maiden <i>et al.</i> , 1998
zv27_519	E26	(group X) R. Moxon / Maiden <i>et al.</i> , 1998
zv28_519	860800	(group Y) R. Moxon / Maiden <i>et al.</i> , 1998
zv29_519ass	E32	(group Z) R. Moxon / Maiden <i>et al.</i> , 1998

**Gonococcus**

zv32_519	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
fa1090_519	FA1090	R. Moxon

**References:**

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
 Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090\_519 <SEQ ID 3185>  
 MEFFIILLAAVAVFGFKSFVVIQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
 KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
 RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE  
 KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
 LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
 ISAGMKIIDSSKTAK\*

Z2491\_519 &lt;SEQ ID 3186&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV01\_519 &lt;SEQ ID 3187&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV02\_519 &lt;SEQ ID 3188&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV03\_519 &lt;SEQ ID 3189&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV04\_519 &lt;SEQ ID 3190&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV05\_519 &lt;SEQ ID 3191&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV06\_519ASS &lt;SEQ ID 3192&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV07\_519 &lt;SEQ ID 3193&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*



ZV11\_519 &lt;SEQ ID 3194&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV12\_519 &lt;SEQ ID 3195&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV18\_519 &lt;SEQ ID 3196&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV19\_519 &lt;SEQ ID 3197&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV20\_519ASS &lt;SEQ ID 3198&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSM  
ISAGMKIIDSSKTAK\*

ZV21\_519ASS &lt;SEQ ID 3199&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV22\_519ASS &lt;SEQ ID 3200&gt;

MEFFIILLVAVVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV26\_519 &lt;SEQ ID 3201&gt;

MEFFIILLVAVVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR  
LVAEANA EAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV27\_519 &lt;SEQ ID 3202&gt;

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV28\_519 <SEQ ID 3203>

MEFFIILLAAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV29\_519ASS <SEQ ID 3204>

MEFFIILLAAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREPEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSNKTAK\*

ZV32\_519 <SEQ ID 3205>

MEFFIILLAAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV96\_519 <SEQ ID 3206>

MEFFIILLAAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQOSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

Figure 22 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 519, further confirming its utility as an antigen for both vaccines and diagnostics.

## EXAMPLE 15

Table 6 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 919 among different strains.

**Table 6**

### 919 gene variability: List of used *Neisseria* strains

Identification Strains	Source / reference
------------------------	--------------------

**number****Group B**

zm01	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm02	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zm03	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm04	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zm05	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zm06	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zm07	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zm08n	528	R. Moxon / Seiler <i>et al.</i> , 1996
zm09	NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zm10	BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zm11asbc	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zm12	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zm13	NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zm14	NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zm15	SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zm16	NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zm17	NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zm18	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zm19	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zm20	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zm21	MC58	R. Moxon
zm96	2996	Our collection

**Group A**

zm22	205900	R. Moxon
zm23asbc	F6124	R. Moxon
z2491	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

**Group C**

zm24	90/18311	R. Moxon
zm25	93/4286	R. Moxon

**Others**

zm26	A22	(group W) R. Moxon / Maiden <i>et al.</i> , 1998
zm27bc	E26	(group X) R. Moxon / Maiden <i>et al.</i> , 1998
zm28	860800	(group Y) R. Moxon / Maiden <i>et al.</i> , 1998
zm29asbc	E32	(group Z) R. Moxon / Maiden <i>et al.</i> , 1998
zm31asbc	<i>N. lactamica</i>	R. Moxon

**Gonococcus**

zm32asbc	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
zm33asbc	Ng SN4	R. Moxon
fa1090	FA1090	R. Moxon

**References:**

- Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3207>

MKKHLLRSALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPAGIPDPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER  
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDGRTERARFPIYGIPTDDFISVPLPAGLRGGKN  
LVRIRQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL  
KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA  
IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

Z2491 <SEQ ID 3208>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER  
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL  
KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELTSNDGPGVAGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM01 <SEQ ID 3209>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL  
KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELTSNDGPGVAGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM02 <SEQ ID 3210>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL  
KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELTSNDGPGVAGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM03 <SEQ ID 3211>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTGTGYEYEPVLKGDGRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL  
KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELTSNDGPGVAGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM04 <SEQ ID 3212>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPLVKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
KLGQTSMQGIKAYMQQNQRRLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM05 <SEQ ID 3213>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPLVKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFREL AGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM06 <SEQ ID 3214>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPLVKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSI GKYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFREL AGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM07 <SEQ ID 3215>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPLVKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFREL AGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM08N <SEQ ID 3216>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPLVKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFREL AGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM09 <SEQ ID 3217>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPLVKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSI GKYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFREL TGSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM10 <SEQ ID 3218>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPLVKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSI GKYMADKGYL

KLQGTSMQGIKSYMRQNPQRLAEVLGQNPYSYIFFRELTGSGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM11ASBC <SEQ ID 3219>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPYSYIFFRELTGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM12 <SEQ ID 3220>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRI RQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPYSYIFFRELAGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM13 <SEQ ID 3221>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRI RQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLQGTSMQGIKAYMRQNPQRLAEVLGQNPYSYIFFRELAGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM14 <SEQ ID 3222>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRI RQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPYSYIFFRELTGSRNDGPVGALGTPLMGEYAGA  
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM15 <SEQ ID 3223>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDLAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNHOGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRI RQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPYSYIFFRELTGSGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM16 <SEQ ID 3224>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPGRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRI RQTGKNSGTIDNTGGHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPYSYIFFRELTGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM17 <SEQ ID 3225>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM18 <SEQ ID 3226>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM19 <SEQ ID 3227>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM20 <SEQ ID 3228>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM21 <SEQ ID 3229>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGQTSMQGIKSYMQRNPORLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM22 <SEQ ID 3230>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGQTSMQGIKAYMQQNORLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM23ASBC <SEQ ID 3231>

MKKYLFRAALYGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEPEVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL

KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELGSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK  
MKEPGYVWQLLPNGMKPEYRP\*

ZM24 <SEQ ID 3232>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM25 <SEQ ID 3233>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM26 <SEQ ID 3234>

MKKYLFRAALYGI AAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKAYMQNPQRLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM27BC <SEQ ID 3235>

MKKYLFRAALYGIS AAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNAGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK  
MKEPGYVWQLLPNGMKPEYRP\*

ZM28 <SEQ ID 3236>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELGSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM29ASBC <SEQ ID 3237>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
LVRIROTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATTHPI TRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM31ASBC <SEQ ID 3238>



MKKHLFRAALYGIAAAILAACQSKSIQTFPQPDTSIIKGPDRPAGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTERARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRI RQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYVFFRELAGSGNDGPVGALGTPLMGEYAGA  
 VDRHYITLGAPL FVATAHPVTRKALNRLIMAQDTGSAIKGAVRVVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM32ASBC <SEQ ID 3239>

MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV  
 YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTERARFPIYGI PDDFISVPLPAGLRGGKA  
 LVRI RQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELAGSGGDPVGALGTPLMGGYAGA  
 IDRHYITLGAPL FVATAHPVTRKALNRLIMAQDTGSAIKGAVRVVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM33ASBC <SEQ ID 3240>

MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV  
 YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTERARFPIYGI PDDFISVPLPAGLRGGKA  
 LVRI RQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELAGSGNEGPGVGALGTPLMGEYAGA  
 IDRHYITLGAPL FVATAHPVTRKALNRLIMAQDTGSAIKGAVRVVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM96 <SEQ ID 3241>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTERARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRI RQTGKNSGTIDNTGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA  
 VDRHYITLGAPL FVATAHPVTRKALNRLIMAQDTGSAIKGAVRVVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

Figure 23 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 919, further confirming its utility as an antigen for both vaccines and diagnostics.

#### EXAMPLE 16

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

**Table 7:** Oligonucleotides used for PCR to amplify complete or partial ORFs

ORF primer Sequence		Restriction sites
001	Forward CGCGGATCCCATATG-TGGATGGTGCTGGTCAT	BamHI-
	Reverse CCCGCTCGAG-TGCCGTCTTGTCCCAC	NdeI
003	Forward CGCGGATCCCATATG-GTCGTATTCGTGGC	XhoI
	Reverse CCCGCTCGAG-AAAATCATGAACACGCGC	BamHI-
005	Forward CGCGGATCCCATATG-GACAATATTGACATGT	NdeI
	Reverse CCCGCTCGAG-CATCACATCCGCCCCG	XhoI
006	Forward CGCGGATCCCATATG-CTGCTGGTGCTGG	BamHI-
	Reverse CCCGCTCGAG-AGTTCCGGCTTTGATGT	NdeI
007	Forward CGCGGATCCCATATG-GCCGACAACAGCATCAT	XhoI
	Reverse CCCGCTCGAG-AAGGCGTTCATGATATAAG	BamHI-
008	Forward CGCGGATCCCATATG-AACAACAGACATTTTG	NdeI
	Reverse CCCGCTCGAG-CCTGTCCGGTAAAAGAC	XhoI
009	Forward CGCGGATCCCATATG-CCCCGCGCTGCT	BamHI-
	Reverse CCCGCTCGAG-TGGCTTTTGCCACGTTTT	NdeI
011	Forward CGCGGATCCCATATG-AAGACACACCGCAAG	XhoI
	Reverse CCCGCTCGAG-GGCGGTCAGTACGGT	BamHI-
012	Forward CGCGGATCCCATATG-CTCGCCCGTTGCC	NdeI
	Reverse CCCGCTCGAG-AGCGGGGAAGAGGCAC	XhoI
013	Forward CGCGGATCCCATATG-CCTTTGACCATGCT	BamHI-
	Reverse CCCGCTCGAG-CTGATTCGGCAAAAAATCT	NdeI
018	Forward CGCGGATCCCATATG-CAGCAGAGGCAGTT	XhoI
	Reverse CCCGCTCGAG-GACGAGGCGAACGCC	BamHI-
019	Forward AAAGAATTC-CTGCCAGCCGGCAAGACCCCGGC	NdeI
	Reverse AAAGTGCAG-TCAGCGGGCGGGGACAATGCCCAT	XhoI
023	Forward AAAGAATTC-AAAGAATATTCGGCATGGCAGGC	Eco RI
	Reverse AAAGTGCAG-TTACCCCCAAATCACTTTAACTGA	Pst I
025	Forward AAAGAATTC-TGCGCCACCCAACAGCCTGCTCC	Eco RI
	Reverse AAAGTGCAG-TCAGAACGCGATATAGCTGTTCCG	Pst I
031	Forward CGCGGATCCCATATG-GTCTCCCTTCGCTT	BamHI-
	Reverse CCCGCTCGAG-ATGTAAGACGGGGACAAC	NdeI
032	Forward CGCGGATCCCATATG-CGGCGAAACGTGC	XhoI
		BamHI-

	Reverse	CCCGCTCGAG-CTGGTTTTTTTGATATTTGTG	NdeI
033	Forward	CGCGGATCCCATATG-GCGGCGGCAGACA	XhoI
	Reverse	CCCGCTCGAG-ATTTGCCGCATCCCGAT	BamHI-
034	Forward	CGCGGATCCCATATG-GCCGAAAACAGCTACGG	NdeI
	Reverse	CCCGCTCGAG-TTTGACGATTTGGTTCAATT	XhoI
036	Forward	CGCGGATCCCATATG-CTGAAGCCGTGCG	BamHI-
	Reverse	CCCGCTCGAG-CCGGACTGCGTATCGG	NdeI
038	Forward	CGCGGATCCCATATG-ACCGATTTCGCCA	XhoI
	Reverse	CCCGCTCGAG-TTCTACGCCGTACTGCC	BamHI-
039	Forward	CGCGGATCCCATATG-CCGTCCGAACCGC	NdeI
	Reverse	CCCGCTCGAG-TAGGATGACGAGGTAGG	XhoI
041	Forward	CGCGGATCCCATATG-TTCGTGCGCGAACCGC	BamHI-
	Reverse	CCCGCTCGAG-GCCCAAAAACCTCTTTCAAA	NdeI
042	Forward	CGCGGATCCCATATG-ACGATGATTTGCTTGC	XhoI
	Reverse	CCCGCTCGAG-TTTGCAGCCTGCATTTGAC	BamHI-
043	Forward	AAAAAAGGTACC-ATGGTTGTTTCAAATCAAATATC	NdeI
	Reverse	AAACTGCAG-TTATTGCGCTTCACCTTCCGCCGC	XhoI
043a	Forward	AAAAAAGGTACC-GCAAAAGTGCATGGCGGCTTGGACGGTGC	Kpn I
	Reverse	AAAAAACTGCAG-TTAATCCTGCAACACGAATTCGCCCCGTCCG	Pst I
044	Forward	CGCGGATCCCATATG-CCGTCCGACTAGAG	Kpn I
	Reverse	CCCGCTCGAG-ATGCGCTACGGTAGCCA	Pst I
046	Forward	AAAGAATTC-ATGTCGGCAATGCTCCCGACAAG	BamHI-
	Reverse	AAACTGCAG-TCACTCGGCGACCCACACCGTGAA	NdeI
047	Forward	CGCGGATCCCATATG-GTCATCATACAGGCG	XhoI
	Reverse	CCCGCTCGAG-TCCGAAAAAGCCCATTTTG	Eco RI
048	Forward	AAAGAATTC-ATGCTCAACAAAGGCGAAGAATTGCC	XhoI
	Reverse	AAACTGCAG-TCAAGATTCGACGGGGATGATGCC	Eco RI
049	Forward	AAAGAATTC-ATGCGGGCGCAGGCGTTTGATCAGCC	Pst I
	Reverse	AAACTGCAG-AAGGCGTATCTGAAAAAATGGCAG	Eco RI
050	Forward	CGCGGATCCCATATG-GGCGCGGGCTGG	Pst I
	Reverse	CCCGCTCGAG-AATCGGGCCATCTTCGA	BamHI-
052	Forward	AAAAAAGAATTC-ATGGCTTTGGTGGCGGAGGAAAC	NdeI
	Reverse	AAAAAAGTGCAG-TCAGGCGGCGTTTTTCACCTTCCT	XhoI
052a	Forward	AAAAAAGAATTC-GTGGCGGAGGAAACGGAAATATCCGC	Eco RI

	Reverse	AAAAAACTGCAG-TTAGCTGTTTTTGGAAACGCCGTCCAACCC	Pst I
073	Forward	CGCGGATCCCATATG-TGTATGCCATATAAGAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CACCGGATTGTCCGAC	XhoI
075	Forward	CGCGGATCCCATATG-CCGTCTTACTTCATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCACCAATGCCGATTATTT	XhoI
077a	Forward	AAAAAAGAATTC-GGCGGCATTTTCATCGACACCTTCCT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGAACATCTGCACAAACGCAAT	Pst I
080	Forward	AAAGAATTC-GCGTCCGGGCTGGTTTGGTTTTACAATTC	Eco RI
	Reverse	AAACTGCAG-CTATTCTTCGGATTCTTTTTCGGG	Pst I
081	Forward	AAAGAATTC-ATGAAACCACTGGACCTAAATTTTCATCTG	Eco RI
	Reverse	AAACTGCAG-TCATTATCCTCCAATGCCTC	Pst I
082	Forward	AAAGAATTC-ATGTGGTTGTTGAAGTTGCCTGC	Eco RI
	Reverse	AAACTGCAG-TTACGCGGATTCGGCAGTTGG	Pst I
084	Forward	AAAGAATTC-TATCACCCAGAATATGAATACGGCTACCG	Eco RI
	Reverse	AAACTGCAG-TTATACTTGGGCGCAACATGA	Pst I
085	Forward	CGCGGATCCCATATG-GGTAAAGGGCAGGACT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CAAAGCCTTAAACGCTTCG	XhoI
086	Forward	AAAAAAGGTACC-TATTTGGCATCAAAAGAAGGCGG	Kpn I
	Reverse	AAACTGCAG-TTACTCCACCCGATAACCGCG	Pst I
087	Forward	AAAGAATTC-ATGGGCGGTAAAACCTTTATGC	Eco RI
	Reverse	AAACTGCAG-TTACGCCGCACACGCAATCGC	Pst I
087a	Forward	AAAAAAGAATTC-AAGCTATTAGGCGTGCCGATTGTGATTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCTGCAAGATGCCAGCTTGCC	Pst I
088	Forward	AAAAAAGAATTC-ATGTTTTTATGGCTCGCACATTTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAGCGGATTTTGAGGGTACTCAAACC	Pst I
089	Forward	CGCGGATCCCATATG-CCGCCCAAATCAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCGCATACCAAAGCCA	XhoI
090	Forward	CGCGGATCCCATATG-CGCATAGTCGAGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCAAAACGGCGGTACG	XhoI
091	Forward	AAAGAATTC-ATGGAAATACCCGTACCGCCGAGTCC	Eco RI
	Reverse	AAACTGCAG-TCAGCGCAGGGGGTAGCCCAAGCC	Pst I
092	Forward	AAAGAATTC-ATGTTTTTTTATTTCAATCCG	Eco RI
	Reverse	AAACTGCAG-TCAAATCTGTTTCGACAATGC	Pst I
093	Forward	AAAGAATTC-ATGCAGAATTTTGGCAAAGTGCC	Eco RI
	Reverse	AAACTGCAG-CTATGGCTCGTCATACCGGGC	Pst I
094	Forward	AAAGAATTC-ATGCCGTCACGGAAGCGCATCAACTC	Eco RI
	Reverse	AAACTGCAG-TTATCCCGGCCATACCGCCGAACA	Pst I
095	Forward	AAAGAATTC-ATGTCCTTTCATTTGAACATGGACGG	Eco RI
	Reverse	AAACTGCAG-TCAACGCCGCAGGCACTAACGCCC	Pst I
096	Forward	AAAGAATTC-ATGGCTCGTCATACCGGGCAGGG	Eco RI

	Reverse	AAACTGCAG-TCAAAGGAAAAGGCCGTCTGAAAAGCG	Pst I
097	Forward	AAAGAATTC-ATGGACACTTCAAAACAAACACTGTTG	Eco RI
	Reverse	AAACTGCAG-TCAGCCCAAATACCAGAATTTTCAG	Pst I
098	Forward	AAAGAATTC-GATGAACGCAGCCAGCATGGATACG	Eco RI
	Reverse	AAACTGCAG-TTACGACATTCTGATTTGGCA	Pst I
102	Forward	AAAAAAGAATTC-GGCCTGATGATTTTGGAAAGTCAACAC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
105	Forward	CGCGGATCCCATATG-TCCGCAAACGAATACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGTTCTGCCAGTTTCAG	XhoI
107	Forward	AAAAAAGAATTC- CTGATGATTTTGGAAAGTCAACACCCATTATCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
107b	Forward	AAAAAAGAATTC- GATACCCAAGCCCCCGCCGGCACAACTACTG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGCGTCGCCTTTAAAGTATTTGAGCAGGCTGGAGAC	Pst I
108	Forward	AAAGAATTC-ATGTTGCCGGGCTTCAACCG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
108a	Forward	AAAAAAGAATTC-GGTAACACATTCGGCAGCTTAGACGGTGG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
109	Forward	AAAGAATTC-ATGTATTATCGCCGGGTTATGGG	Eco RI
	Reverse	AAACTGCAG-CTAGCCCAAAGATTTGAAGTGTTT	Pst I
111	Forward	CGCGGATCCCATATG-TGTTTCGGAACAAACCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGAGCAGTTTTTCAA	XhoI
114	Forward	CGCGGATCCCATATG-GCTTCCATCACTTCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATCCGCGAAATCGTC	XhoI
117	Forward	AAAAAAGGTACC-ATGGTCAAGAACTGGAAGTCTG	Kpn I
	Reverse	AAACTGCAG-TTAAAGCCGGGTAACGCTCAATAC	Pst I
118	Forward	AAAGTCGACATGTGTGAGTTCAAGGATATTATAAG	Sal I
	Reverse	AAAGCATGC-CTATTTTTTGTGTGAATAATCAAATC	Sph I
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC	XhoI
122	Forward	CGCGGATCCCATATG-GTCATGATTAAATCCGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AATCTTGGTAGATTGGATT	XhoI
125	Forward	AAAGAATTC-ATGTCGGGCAATGCCTCCTCTCC	Eco RI
	Reverse	AAACTGCAG-TCACGCCGTTTCAAGACG	Pst I
125a	Forward	AAAAAAGAATTC-ACGGCAGGCAGCACCGCCGCACAGGTTTC	Eco RI
	Reverse	AAAAAACTGCAG- TTATTTTGGCACGTCGGTTTCTCCGGTGAACAACGC	Pst I
126	Forward	CGCGGATCCCATATG-CCGTCTGAAACCC	BamHI-

	Reverse	CCCGCTCGAG-ATATTCCGCCGAATGCC	NdeI
127	Forward	AAAGAATTC-ATGGAAATATGGAATATGTTGGACACTTG	XhoI
	Reverse	AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Eco RI
			Pst I
127a	Forward	AAAAAAGAATTC-AAGGAACTGATTATGTGTCTGTCTGGG	Eco RI
	Reverse	AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Pst I
128	Forward	CGCGGATCCCATATG-ACTGACAACGCACT	BamHI-
	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA	NdeI
130	Forward	CGCGGATCCCATATG-AAACAACTCCGCGA	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-GAATTTTGCACCGGATTG	NdeI
132	Forward	AAAGAATTC-ATGGAACCCTTCAAAACCTTAATTTG	XhoI
	Reverse	AAAAAACTGCAG-TCACCATGTCTGGCATTGAAAAAC	Eco RI
134	Forward	CGCGGATCCCATATG-TCCCAAGAAATCCTC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CAGTTTGACCGAATGTTC	NdeI
135	Forward	CGCGGATCCCATATG-AAATACAAAAGAATCGTATT	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-AAATTCGGTCAGAAGCAGG	NdeI
137	Forward	AAAAAAGGTACC-ATGATTACCCATCCCCAATTCGATCC	XhoI
	Reverse	AAAAAACTGCAG-TCAGTGCTGTTTTTTCATGCCGAA	Kpn I
137a	Forward	AAAAAAGAATTC-GGCCGCAAACACGGCATCGGCTTCCT	Pst I
	Reverse	AAAAAACTGCAG-TTAAGCGGGATGACGCGGCAGCATACC	Eco RI
138	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Pst I
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Eco RI
141	Forward	AAAGAATTC-ATGAGCTTCAAAACCGATGCCGAAATCGC	Xba I
	Reverse	AAACTGCAG-TCAGAACAAGCCGTGAATCACGCC	Eco RI
142	Forward	CGCGGATCCCATATG-CGTGCCGATTTTCATG	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-AAACTGCTGCACATGGG	NdeI
143	Forward	AAAAAAGAATTC- ATGCTCAGTTTCGGCTTTCTCGGCGTTCAGAC	XhoI
	Reverse	AAAAAACTGCAG-TCAAACCCCGCCGTGTGTTTCTTTAAT	Eco RI
144	Forward	AAAAAAGAATTC-GGTCTGATCGACGGGCGTGCCGTAAC	Pst I
	Reverse	AAAAAATCTAGA-TCGGCATCGGCCGGCATATGTCCG	Eco RI
146	Forward	AAAAAAGAATTC- CGCCAAGTCGTCATTGACCACGACAAAGTC	Xba I
	Reverse	AAAAAACTGCAG-TTAGGCATCGGCAAATAGGAACTGGG	Eco RI
147	Forward	AAAAAAGAATTC-ACTGAGCAATCGGTGGATTGGAAC	Pst I
	Reverse	AAAAAATCTAGA-TTAGGTAAAGCTGCGGCCCATTTGCGG	Eco RI
148	Forward	AAAAAAGAATTC- ATGGCGTTAAAAACATCAAACCTTGAACACGC	Xba I
	Reverse	AAAAAATCTAGA-TCAGCCCTTCATACAGCCTTCGTTTTG	Eco RI
149	Forward	CGCGGATCCCATATG-CTGCTTGACAACAAAGT	Xba I
			BamHI-

	Reverse	CCCGCTCGAG-AAACTTCACGTTACGCC	NdeI
150	Forward	CGCGGATCCCATATG-CAGAACACAAATCCG	XhoI
	Reverse	CCCGCTCGAG-ATAAACATCACGCTGATAGC	BamHI-
151	Forward	AAAAAAGAATTC- ATGAAACAAATCCGCAACATCGCCATCATCGC	NdeI
	Reverse	AAAAAACTGCAG-TCAATCCAGCTTTTAAAGTGGCGGCG	XhoI
152	Forward	AAAAAAGAATTC- ATGAAAAACAAAACCAAAGTCTGGGACCTCCC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGGACAGGAGCAGGATGGCGGC	Pst I
153	Forward	AAAAAAGAATTC-ATGGCGTTTGCTTACGGTATGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTCATGTTTTTCCGTTTCATT	Pst I
153a	Forward	AAAAAAGAATTC-CGGACTTCGGTATCGGTTCCCAGCATTG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGCCGACGAAATACTCAGACTTTTCGG	Pst I
154	Forward	CGCGGATCCCATATG-ACTGACAACAGCCC	BamHI-
	Reverse	CCCGCTCGAG-TCGGCTTCCTTTTCGGG	NdeI
155	Forward	AAAAAAGAATTC-ATGAAAATCGGTATCCCACGCGAGTC	XhoI
	Reverse	AAAAAACTGCAG-TTACCCTTTCTTAAACATATTCAGCAT	Eco RI
156	Forward	AAAAAAGAATTC-GCACAGCAAAACGGTTTTGAAGC	Pst I
	Reverse	AAAAAACTGCAG-TCAAGCAGCCGCGACAAACAGCCC	Eco RI
157	Forward	CGCGGATCCCATATG-AGGAACGAGGAAAAAC	Pst I
	Reverse	CCCGCTCGAG-AAAACACAATATCCCCGC	BamHI-
158	Forward	AAAAAAGAATTC-GCGGAGCAGTTGGCGATGGCAAATTCTGC	NdeI
	Reverse	AAAAAATCTAGA-TTATCCACAGAGATTGTTTCCCAGTTC	XhoI
160	Forward	CGCGGATCCCATATG-GACATTCTGGACAAAC	Eco RI
	Reverse	CCCGCTCGAG-TTTTTGCCCGCCTTCTTT	Xba I
163	Forward	AAAAAAGGTACC-ACCGTGCCGGATCAGGTGCAGATGTG	Kpn I
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Xba I
163a	Forward	AAAAAAGAATTC-CGGCTGGTGCAGATAATGAGCCAGAC	Eco RI
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Xba I
164	Forward	CGCGGATCCCATATG-AACCGGACTTATGCC	BamHI-
	Reverse	CCCGCTCGAG-TTTGTTTCCGTCAAACCTGC	NdeI
165	Forward	CGCGGATCCGCTAGC-GCTGAAGCGACAGACG	XhoI
	Reverse	CCCGCTCGAG-AATATCCAATACTTTTCGCG	BamHI-
206	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA	NheI
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC	XhoI
209	Forward	CGCGGATCCCATATG-CTGCGGCATTTAGGA	BamHI-
			NdeI

	Reverse	CCCGCTCGAG-TACCCCTGAAGGCAAC	XhoI
211	Forward	AAAAAAGAATTC-ATGTTGCGGGTTGCTGCTGC	Eco RI
	Reverse	AAAAAACTGCAG-CTATCCTGCGGATTGGCATTGAAA	Pst I
212	Forward	CGCGGATCCCATATG-GACAATCTCGTATGG	BamHI-
	Reverse	CCCGCTCGAG-AGGGGTAGATCCTTCC	NdeI
215	Forward	CGCGGATCCCATATG-GCATGGTTGGGTCGT	XhoI
	Reverse	CCCGCTCGAG-CATATCTTTTGTATCATAAATC	BamHI-
216	Forward	CGCGGATCCCATATG-GCAATGGCAGAAAACG	NdeI
	Reverse	CCCGCTCGAG-TACAATCCGTGCCGCC	XhoI
217	Forward	CGCGGATCCCATATG-GCGGATGACGGTGTG	BamHI-
	Reverse	CCCGCTCGAG-ACCCCGAATATCGAATCC	NdeI
218	Forward	CGCGGATCCCATATG-GTCGCGGTCGATC	XhoI
	Reverse	CCCGCTCGAG-TAACTCATAGAATCCTGC	BamHI-
219	Forward	CGCGGATCCGCTAGC-ACGGCAAGGTAAAG	NheI
	Reverse	CCCGCTCGAG-TTTAAACCATCTCCTCAAAAC	XhoI
223	Forward	CGCGGATCCCATATG-GAATTCAGGCACCAAGTA	BamHI-
	Reverse	CCCGCTCGAG-GGCTTCCCGCGTGTG	NdeI
225	Forward	CGCGGATCCCATATG-GACGAGTTGACCAACC	XhoI
	Reverse	CCCGCTCGAG-GTTCAGAAAGCGGGAC	BamHI-
226	Forward	AAAGAATTC-CTTGCGATTATCGTGCGCACGCG	NdeI
	Reverse	AAACTGCAG-TCAAAATCCCAAACGGGGAT	XhoI
228	Forward	CGCGGATCCCATATG-TCGCAAGAAGCCAAACAG	Eco RI
	Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	Pst I
229	Forward	CGCGGATCCCATATG-CAAGAGGTTTTGCC	BamHI-
	Reverse	CCCGCTCGAG-ACACAATATAGCGGATGAAC	NdeI
230	Forward	CGCGGATCCCATATG-CATCCGGGTGCCGAC	XhoI
	Reverse	CCCGCTCGAG-AAGTTTGGCGGCTTCGG	BamHI-
232	Forward	AAAAAAGAATTC-ATGTACGCTAAAAAAGGCGGTTTGGG	NdeI
	Reverse	AAAAAACTGCAG-TCAAGGTTTTTTCCTGATTGCCGCCGC	XhoI
232a	Forward	AAAAAAGAATTC-GCCAAGGCTGCCGATACACAAATTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACATTGTCGTTGCCGCCAGATG	Pst I
233	Forward	CGCGGATCCCATATG-GCGGACAAACCCAAG	BamHI-
	Reverse	CCCGCTCGAG-GACGGCATTGAGCAG	NdeI
234	Forward	CGCGGATCCCATATG-GCCGTTTCACTGACCG	XhoI
			BamHI-



	Reverse	GCCCAAGCTT-ACGGTTGGATTGCCATG	NdeI
235	Forward	CGCGGATCCCATATG-GCCTGCCAAGTTCAAA	Hind III
	Reverse	CCCGCTCGAG-TTTGGGCTGCTCTTC	BamHI-
236	Forward	CGCGGATCCCATATG-GCGCGTTTCGCCTT	NdeI
	Reverse	CCCGCTCGAG-ATGGGTCGCGCGCCGT	XhoI
238	Forward	CGCGGATCCGCTAGC-AACGGTTTGGATGCCCCG	BamHI-
	Reverse	CCCGCTCGAG-TTTGTCTAAGTTCCTGATATG	NdeI
239	Forward	CCGGAATTCTACATATG-CTCCACCATAAAGGTATTG	XhoI
	Reverse	CCCGCTCGAG-TGGTGAAGAGCGGTTTAG	EcoRI-
240	Forward	CGCGGATCCCATATG-GACGTTGGACGATTTC	NdeI
	Reverse	CCCGCTCGAG-AAACGCCATTACCCGATG	XhoI
241	Forward	CCGGAATTCTACATATG-CCAACACGTCCAAC	BamHI-
	Reverse	CCCGCTCGAG-GAATGCGCCTGTAATTAATC	NdeI
242	Forward	CGCGGATCCCATATG-ATCGGCAAACCTTGTTG	XhoI
	Reverse	GCCCAAGCTT-ACCGATACGGTCGCAG	BamHI-
243	Forward	CGCGGATCCCATATG-ACGATTTTTTCGATGCTGC	NdeI
	Reverse	CCCGCTCGAG-CGACTTGGTTACCGCG	XhoI
244	Forward	CGCGGATCCCATATG-CCGTCTGAAGCCC	BamHI-
	Reverse	CCCGCTCGAG-TTTTTTCGGTAGGGGATTT	NdeI
246	Forward	CGCGGATCCCATATG-GACATCGGCAGTGC	XhoI
	Reverse	CCCGCTCGAG-CCCGCGCTGCTGGAG	BamHI-
247	Forward	CGCGGATCCCATATG-GTCGGATCGAGTTAC	NdeI
	Reverse	CCCGCTCGAG-AAGTGTTCTGTTTGCGCA	XhoI
248	Forward	CGCGGATCCCATATG-CGCAAACAGAACACT	BamHI-
	Reverse	CCCGCTCGAG-CTCATCATTATTGCTAACA	NdeI
249	Forward	CGCGGATCCCATATG-AAGAATAATGATTGCTTC	XhoI
	Reverse	CCCGCTCGAG-TTCCCGACCTCCGAC	BamHI-
251	Forward	CGCGGATCCCATATG-CGTGCTGCGGTAGT	NdeI
	Reverse	CCCGCTCGAG-TACGAAAGCCGGTCGTG	XhoI
253	Forward	AAAAAAGAATTC-ATGATTGACAGGAACCGTATGCTGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I

253a	Forward	AAAAAAGAATTC-AAAATCCTTTTGAAAACAAGCGAAAACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I
254	Forward	AAAAAAGAATTC-ATGTATACAGGCGAACGCTTCAATAC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATTACGTAACCGTACACGCTGAC	Xba I
255	Forward	CGCGGATCCCATATG-GCCGCGTTGCGTTAC	BamHI-
	Reverse	CCCGCTCGAG-ATCCGCAATACCGACCAG	NdeI
256	Forward	CGCGGATCCGCTAGC-TTTTAACACCGCCGGAC	XhoI
	Reverse	CCCGCTCGAG-ACGCCTGTTTGTGCGG	BamHI-
257	Forward	CGCGGATCCCATATG-GCGGTTTCTTTCCTG	NheI
	Reverse	CCCGCTCGAG-GCGCGTGAATATCGCG	XhoI
258	Forward	AAAAAAGAATTC-GATTATTTCTGGTGGATTGTTGCGTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
258a	Forward	AAAAAAGAATTC-GCGAAGGCGGTGGCGCAAGGCGA	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
259	Forward	CGCGGATCCCATATG-GAAGAGCTGCCTCCG	BamHI-
	Reverse	CCCGCTCGAG-GGCTTTTCCGGCGTTT	NdeI
260	Forward	CGCGGATCCCATATG-GGTGCGGGTATGGT	XhoI
	Reverse	CCCGCTCGAG-AACAGGGCGACACCCT	BamHI-
261	Forward	AAAAAAGAATTC-CAAGATACAGCTCGGGCATTTCGC	NdeI
	Reverse	AAAAAACTGCAG-TCAAACCAACAAGCCTTGGTCACT	XhoI
263	Forward	CGCGGATCCCATATG-GCACGTTTAACCGTA	Eco RI
	Reverse	CCCGCTCGAG-GGCGTAAGCCTGCAATT	Pst I
264	Forward	AAAAAAGGTACC-GCCGACGCAAGTGGTCAAGGCAGAA	Eco RI
	Reverse	AAACTGCAG-TCAGCCGGCGGTCAATACCGCCCG	Pst I
265	Forward	AAAAAAGAATTC-GCGGAGGTCAAGAGAAGGTGTTTG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGAATACGTCGTCAAATGGG	Pst I
266	Forward	AAAGAATTC-CTCATCTTTGCCAACGCCCCCTTC	Eco RI
	Reverse	AAACTGCAG-CTATTCCCTGTTGCGCGTGTGCCA	Pst I
267	Forward	AAAGAATTC-TTCTTCCGATTTCGATGTTAATCG	Eco RI
	Reverse	AAACTGCAG-TTAGTAAAAACCTTTCTGCTTGGC	Pst I
269	Forward	AAAGAATTC-TGCAAACCTTGCGCCACGTGCCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
269a	Forward	AAAAAAGAATTC-GACTTTATCCAAAACACGGCTTCGCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
270	Forward	AAAGAATTC-GCCGTCAAGCTCGTTTTGTGCAATG	Eco RI
	Reverse	AAACTGCAG-TTATTCGGCGGTAAATGCCGTCTG	Pst I
271	Forward	CGCGGATCCCATATG-CCTGTGTGCAGCTCGAC	BamHI-
	Reverse	CCCGCTCGAG-TCCCAGCCCCGTGGAG	NdeI
			XhoI

272	Forward	<u>AAAGAATTC</u> -ATGACCGCAAAGGAAGAACTGTTTCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAGCAGTTCCAAATCGGGGCT	Pst I
273	Forward	<u>AAAGAATTC</u> -ATGAGTCTTCAGGCGGTATTTATATACCC	Eco RI
	Reverse	<u>AAACTGCAG</u> -TTACGCGTAAGAAAAAACTGC	Pst I
274	Forward	<u>CGCGGATCCCATATG</u> -ACAGATTTGGTTACGGAC	BamHI- NdeI
	Reverse	<u>CCCGCTCGAG</u> -TTTGCTTTCAGTATTATTGAA	XhoI
276	Forward	<u>AAAAAAGAATTC</u> - ATGATTTTGCCGTCGTCCATCACGATGATGCG	Eco RI
	Reverse	<u>AAAAAACTGCAG</u> -CTACACCACCATCGGCGAATTTATGGC	Pst I
277	Forward	<u>AAAAAAGAATTC</u> -ATGCCCCGCTTTGAGGACAAGCTCGTAGG	Eco RI
	Reverse	<u>AAAAAACTGCAG</u> -TCATAAGCCATGCTTACCTTCCAACAA	Pst I
277a	Forward	<u>AAAAAAGAATTC</u> -GGGGCGGCGGCTGGGTGACGTAGG	Eco RI
	Reverse	<u>AAAAAACTGCAG</u> -TCATAAGCCATGCTTACCTTCCAACAA	Pst I
278	Forward	<u>AAAAAAGGTACC</u> -GTCAAAGTTGTATTAATCGGGCCTTTGCC	Kpn I
	Reverse	<u>AAAAAACTGCAG</u> -TCATTCAACCATATCAAATCTGCC	Pst I
278a	Forward	<u>AAAAAAGAATTC</u> -AAAACCTCCTAATTCGTCATAGTCG	Eco RI
	Reverse	<u>AAAAAACTGCAG</u> -TCATTCAACCATATCAAATCTGCC	Pst I
279	Forward	<u>CGCGGATCCCATATG</u> -TTGCCTGCAATCACGATT	BamHI- NdeI
	Reverse	<u>CCCGCTCGAG</u> -TTTAGAAGCGGGCGGCAA	XhoI
280	Forward	<u>AAAAAAGGTACC</u> -GCCCCCCTGCCGTTGTAACCAG	Kpn I
	Reverse	<u>AAAAAACTGCAG</u> -TTATTGCTTCATCGCGTTGGTCAAGGC	Pst I
281	Forward	<u>AAAAAAGAATTC</u> -GCACCCGTCGGCGTATTCCTCGTCATGCG	Eco RI
	Reverse	<u>AAAAAATCTAGA</u> -GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
281a	Forward	<u>AAAAAAGAATTC</u> -TCCTACCACATCGAAATTCCTTCCGG	Eco RI
	Reverse	<u>AAAAAATCTAGA</u> -GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
282	Forward	<u>AAAAAAGAATTC</u> -CTTTACCTTGACCTGACCAACGGGCACAG	Eco RI
	Reverse	<u>AAAAAACTGCAG</u> -TCAACCTGCCAGTTGCGGGAATATCGT	Pst I
283	Forward	<u>CGCGGATCCCATATG</u> -GCCGTCTTTACTTGGAAG	BamHI- NdeI
	Reverse	<u>CCCGCTCGAG</u> -ACGGCAGTATTTGTTTACG	XhoI
284	Forward	<u>CGCGGATCCCATATG</u> -TTTGCCTGCAAAGAATCG	BamHI- NdeI
	Reverse	<u>CCCGCTCGAG</u> -CCGACTTTGCAAAAACCTG	XhoI
286	Forward	<u>CGCGGATCCCATATG</u> -GCCGACCTTTCCGAAAA	BamHI- NdeI
	Reverse	<u>CCCGCTCGAG</u> -GAAGCGCGTTCCCAAG	XhoI
287	Forward	<u>CCGGAATTCTAGCTAGC</u> -CTTTCAGCCTGCGGG	EcoRI- NheI
	Reverse	<u>CCCGCTCGAG</u> -ATCCTGCTCTTTTTTTGCC	XhoI
288	Forward	<u>CGCGGATCCCATATG</u> -CACACCGGACAGG	BamHI- NdeI
	Reverse	<u>CCCGCTCGAG</u> -CGTATCAAAGACTTGCGT	XhoI
290	Forward	<u>CGCGGATCCCATATG</u> -GCGGTTTGGGGCGGA	BamHI-

	Reverse	CCCGCTCGAG-TCGGCGCGGCGGGC	NdeI
292	Forward	CGCGGATCCCATATG-TGCGGGCAAACGCCC	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-TTGATTTTTCGGGATGATT	NdeI
294	Forward	AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC	XhoI
	Reverse	AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	Eco RI
295	Forward	CGCGGATCCCATATG-AACCGGCCGCGCCTCC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CGATATTTGATTCCGTTGC	NdeI
297	Forward	AAAAAAGAATTC-GCATAATTGCTTCGACAGAGAG	XhoI
	Reverse	AAAAAACTGCAG-TCAATCCGATTGCGACACGGT	Eco RI
298	Forward	AAAAAAGAATTC-CTGATTGCCGTGTGGTTCAGCCAAAACCC	Pst I
	Reverse	AAAAAACTGCAG-TCATGGCTGTGTAATTGATGGTTGCGT	Eco RI
299	Forward	CGCGGATCCGCTAGC-CTACCTGTCGCCTCCG	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-TTGCCTGATTGCAGCGG	NheI
302	Forward	AAAAAAGAATTC-ATGAGTCAAACCGATACGCAACG	XhoI
	Reverse	AAAAAACTGCAG-TTAAGGTGCGGGATAGAATGTGGGGCGC	Eco RI
305	Forward	AAAAAAGGTACC-GAATTTTACCGATTTCAGCACCGGA	Pst I
	Reverse	AAAAAACTGCAG-TCATTCCCACTTATCCAGCCTGACAG	Kpn I
305a	Forward	AAAAAAGGTACC-TCCCGTTCGGGCAGTACGATTATGGG	Pst I
	Reverse	AAAAAACTGCAG-TTACAAACCGACATCATGCAGGGTGAA	Kpn I
306	Forward	CGCGGATCCCATATG-TTTATGAACAAATTTTCCC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CCGCATCGGCAGAC	NdeI
308	Forward	CGCGGATCCCATATG-TTAAATCGGGTATTTTATC	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-ATCCGCCATTCCCTGC	NdeI
311	Forward	AAAAAAGGTACC-ATGTTTCAGTTTTGGCTGGGTGTTT	XhoI
	Reverse	AAACTGCAG-ATGTTTCATATTCCCTGCCTTCGGC	Kpn I
312	Forward	AAAAAAGGTACC-ATGAGTATCCCATCCGGCGAAATT	Pst I
	Reverse	AAACTGCAG-TCAGTTTTTCATCGATTGAACCGG	Kpn I
313	Forward	AAAAAAGAATTC-ATGGACGACCCGCGCACCTACGGATC	Pst I
	Reverse	AAAAAACTGCAG-TCAGCGGCTGCCGCCGATTTTGCT	Eco RI
401	Forward	CGCGGATCCCATATG-AAGGCGGCAACACAGC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CCTTACGTTTTTCAAAGCC	NdeI
402	Forward	AAAAAAGAATTC-GTGCCTCAGGCATTTTCATTTACCCTTGC	XhoI
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Eco RI
402a	Forward	AAAAAAGAATTC-AGGCTGATTGAAAACAAACACGG	Xba I
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Eco RI
406	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG	Xba I
			BamHI-
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG	NdeI
			XhoI

501	Forward	CGCGGATCCCATATG-GCAGGCGGAGATGGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGTGTGATGTTACCC	XhoI
502	Forward	CGCGGATCCCATATG-GTAGACGCGCTTAAGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCTGCATGGCGGCG	XhoI
503	Forward	CGCGGATCCCATATG-TGTTCGGGGAAAGGCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCGCGCATTCTCGCA	XhoI
504	Forward	CGCGGATCCCATATG-AGCGATATTGAAGTGACG	BamHI- NdeI
	Reverse	GCCCAAGCTT-TGATTCAAGTCCTTGCCG	HindIII
505	Forward	CGCGGATCCCATATG-TTTCGTTTACAATTCAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCGTTTTATAGCGG	XhoI
510	Forward	CGCGGATCCCATATG-CCTTCGCGGACAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCACTGGCAGCG	XhoI
512	Forward	CGCGGATCCCATATG-GGACATGAAGTAACGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGAATAGCCTTTGACG	XhoI
515	Forward	CGCGGATCCCATATG-GAGGAAATAGCCTTCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATGCCGCAAAGCATC	XhoI
516	Forward	CGCGGATCCCATATG-TGTACGTTGATGTTGTGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGCGGGCGGCATC	XhoI
517	Forward	CGCGGATCCCATATG-GGTAAAGGTGTGGAAATA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGCGCCAGCCGT	XhoI
518	Forward	AAAGAATTC-GCTTTTTTACTGCTCCGACCGGAAGG	Eco RI
	Reverse	AAACTGCAG-TCAAATTTACAGACTCTGCCAC	Pst I
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC	XhoI
520	Forward	CGCGGATCCCATATG-CCTGCGCTTCTTTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATATTTACATTTAGTCGGC	XhoI
521	Forward	CGCGGATCCCATATG-GCCAAAATCTATACCTGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATACGCCCCAGTTCC	XhoI
522	Forward	CGCGGATCCCATATG-ACTGAGCCGAAACAC	BamHI- NdeI
	Reverse	GCCCAAGCTT-TTCTGATTTCAAATCGGCA	HindIII
523	Forward	CGCGGATCCCATATG-GCTCTGCTTTCCGCG	BamHI- NdeI

	Reverse	CCCGCTCGAG-AGGGTGTGTGATAATAAGAAG	XhoI
525	Forward	CGCGGATCCCATATG-GCCGAAATGGTTCAAATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCCGTGCATATCATAAA	XhoI
527	Forward	AAAGAATTC-TTCCCTCAATGTTGCCGTTTTCG	Eco RI
	Reverse	AAACTGCAG-TTATGCTAAACTCGAAACAAATTC	Pst I
529	Forward	CGCGGATCCGCTAGC-TGCTCCGGCAGCAAAAC	BamHI- NheI
	Reverse	GCCCAAGCTT-ACGCAGTTCGGAATGGAG	HindIII
530	Forward	CGCGGATCCCATATG-AGTGCGAGCGCGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGACCGACTGATTCCG	XhoI
531	Forward	AAAAAAGAATTC-TATGCCGCCGCCTACCAAATCTACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACAGCGCCGTGCCGACGACAAG	Pst I
532	Forward	AAAAAAGAATTC-ATGAGCGGTCAGTTGGGCAAAGGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCCAAGTGGTCGGTATCAAA	Pst I
532a	Forward	AAAAAAGAATTC-TTGGGTGTCGCGTTTGGAGCCGGAAGT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCCAAGTGGTCGGTATCAAA	Pst I
535	Forward	AAAGAATTC-ATGCCCTTTCCCGTTTTCAGAC	Eco RI
	Reverse	AAACTGCAG-TCAGACGACCCCGCCTTCCCC	Pst I
537	Forward	CGCGGATCCCATATG-CATACCCAAAACCAATCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCCTGCAAATAAAGGGTT	XhoI
538	Forward	CGCGGATCCCATATG-GTCGAGCTGGTCAAAGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGGCATTTCGGTTTCGTC	XhoI
539	Forward	CGCGGATCCGCTAGC-GAGGATTTGCAGGAAA	BamHI- NheI
	Reverse	CCCGCTCGAG-TACCAATGTCGGCAAATC	XhoI
542	Forward	AAAGAATTC-ATGCCGTCTGAAACCGTGTC	Eco RI
	Reverse	AAACTGCAG-TTACCGCGAACCGGTCAGGAT	Pst I
543	Forward	AAAAAAGAATTC-GCCTTCGATGGCGACGTTGTAGGTAC	Eco RI
	Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
543a	Forward	AAAAAAGAATTC-GGCAAAACTCGTCATGAATTTGC	Eco RI
	Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
544	Forward	AAAGAATTC-GCGCCCGCCTTCTCCCTGCCCGACCTGCACGG	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
544a	Forward	AAAAAAGAATTC- GCAATGACTATAAAAAACAAAACTTCCAAGTACTTGC	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
547	Forward	AAAGAATTC-ATGTTCTGTAGATAACGGATTTAATAAAAC	Eco RI
	Reverse	AAACTGCAG-TTAACAACAAAAACAAACCGCTT	Pst I
548	Forward	AAAGAATTC-GCCTGCAAACCTCAAGACAACAGTGCGGC	Eco RI

	Reverse	AAACTGCAG-TCAGAGCAGGGTCCTTACATCGGC	Pst I
550	Forward	AAAAAAGTCGAC- ATGATAACGGACAGGTTTCATCTCTTTTCATTTTCC	Sal I
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
550a	Forward	AAAAAAGAATTC-GTAAATCACGCCTTTGGAGTCGCAAACGG	Eco RI
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
552	Forward	AAAAAAGAATTC-TTGGCGCGTTGGCTGGATAC	Eco RI
	Reverse	AAACTGCAG-TTATTTCTGATGCCTTTTCCCAAC	Pst I
554	Forward	CGCGGATCCCATATG-TCGCCCGCGCCCAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGCCCTGTCAGACAC	XhoI
556	Forward	AAAGAATTC-GCGGGCGGTTTTGTTTGGACATCCCCG	Eco RI
	Reverse	AAACTGCAG-TTAACGGTGCGGACGTTTCTGACC	Pst I
557	Forward	CGCGGATCCCATATG-TGCGGTTTCCACCTGAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTCCGCCTTCAGAAAGG	XhoI
558	Forward	AAAGAATTC-GAGCTTTATATGTTTCAACAGGGGACGGC	Eco RI
	Reverse	AAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
558a	Forward	AAAAAAGAATTC-ATTAGATTCTATCGCCATAAACAGACGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
560	Forward	AAAAAAGAATTC- TCGCCTTTCCGGGACGGGGCGCACAAGATGGC	Eco RI
	Reverse	AAAAAACTGCAG-TCATGCGGTTTTCAGACGGCATTTTGGC	Pst I
561	Forward	CCGGAATTCTACATATG-ATACTGCCAGCCCGT	EcoRI- NdeI
	Reverse	CCCGCTCGAG-TTTCAAGCTTTCTTCAGATG	XhoI
562	Forward	CGCGGATCCCATATG-GCAAGCCCGTCGAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGACCAACTCCAACCTCGT	XhoI
565	Forward	CGCGGATCCCATATG-AAGTCGAGCGCGAAATAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCATTGATCGGCGGC	XhoI
566	Forward	CGCGGATCCCATATG-GTCGGTGGCGAAGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGCATGGGCGAAGTCA	XhoI
567	Forward	CCGGAATTCTACATATG-AGTGCGAACATCCTTG	EcoRI- NdeI
	Reverse	CCCGCTCGAG-TTTCCCCGACACCCTCG	XhoI
568	Forward	CGCGGATCCCATATG-CTCAGGGTCAGACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCGCGGCGTTTCAG	XhoI
569	Forward	AAAAAAGAATTC-CTGATTGCCTTGTGGGAATATGCCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATGCATAGACGCTGATAACGGCAAT	Pst I
570	Forward	CGCGGATCCCATATG-GACACCTTCCAAAAAATCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGGCGTTTCATTTCTTT	XhoI

571	Forward	AAAAAAGAATTC- ATGGGTATTGCCGGCGCCGTAAATGTTTTGAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATGGCCGACGCGCGGCTACCTGACG	Pst I
572	Forward	CGCGGATCCCATATG-GCGCAAAAAGGCAAAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCAGTGTGCCGATA	XhoI
573	Forward	CGCGGATCCCATATG-CCCTGTTTGTGCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGGTGTCAATTCGCC	XhoI
574	Forward	CGCGGATCCCATATG-TGGTTTGCCGCCCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACTTCGATTTTATTCGGG	XhoI
575	Forward	CGCGGATCCCATATG-GTTTCGGGCGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATTCCGAATCTGAACAG	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC	XhoI
577	Forward	CGCGGATCCCATATG-GAAAGGAACGGTGTATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGCTGTTTGGTAGATTCTG	XhoI
578	Forward	CGCGGATCCCATATG-AGAAGGTTCGTACAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCAACGCCTCCACG	XhoI
579	Forward	CGCGGATCCCATATG-AGATTGGGCGTTTCCAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGAATTGATGATGTGTATGT	XhoI
580	Forward	CGCGGATCCCATATG-AGGCAGACTTCGCCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-CACTTCCCCCGAAGTG	XhoI
581	Forward	CGCGGATCCCATATG-CACTTCGCCCAGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGCCGTTTGGCTTTGG	XhoI
582	Forward	AAAAAAGAATTC-TTTGGAGAGACCGCGCTGCAATGCGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATGCCGTCCAGTCGTTGAA	Xba I
583	Forward	AAAAAAGAATTC-ACTGCCGGCAATCGACTGCATAATCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGAGGTCAATATGATGAAATTG	Pst I
584	Forward	AAAAAAGAATTC- GCGGCTGAAGCATTGAATTACAATATTGTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGAACTGAACCGTCCCATGACGCT	Pst I
585	Forward	AAAAAAGGTACC-TCTTTCTGGCTGGTGCAGAACACCCTTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTTCGCACTTTTTTCTGTTTTGGA	Pst I
586	Forward	CGCGGATCCCATATG-GCAGCCCATCTCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTCAGCGAATCAAGTTTC	XhoI
587	Forward	CGCGGATCCCATATG-GACCTGCCCTTGACGA	BamHI-



	Reverse	CCCGCTCGAG-AAATGTATGCTGTACGCC	NdeI
			XhoI
588	Forward	AAAAAAGAATTC-GCCGTCCTGACTTCCTATCAAGAACCAGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTGTTTTTGGGCAGTTTCACTTC	Pst I
589	Forward	AAAAAAGAATTC- ATGCAACAAAAAATCCGTTTCCAAATCGAAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAATCGATTTTTACCCGTTTCAGGCG	Pst I
590	Forward	AAAAAAGAATTC-ATGAAAAAACCTTTGATTTCAAGTTGCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCTGCGGCTCTGAAACCAT	Pst I
591	Forward	AAAAAAGAATTC-CACTACATCGTTGCCAGATTGTGCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAACCGAGCAGCCGGGTAACGTCGTT	Pst I
592a	Forward	AAAAAAGAATTC-CGCGATTACACCGCCAAGCTGAAAATGGG	Eco RI
	Reverse	AAAAAACTGCAG-TTACCAAACGTCGGATTTGATACG	Pst I
593	Forward	CGCGGATCCGCTAGC-CTTGAACGTAACGGACTC	BamHI- NheI
	Reverse	CCCGCTCGAG-GCGGAAGCGGACGATT	XhoI
594a	Forward	AAAAAAGAATTC-GGTAAGTTCGCCGTTTCAGGCCTTTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCGCGGTTTCTGACACTCGCG	Pst I
595	Forward	AAAAAAGAATTC-TGCCAGCCGCGGAGGCGGAGAAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCAAGCCGAGTATGCCGCG	Pst I
596	Forward	CGCGGATCCCATATG-TCCCAACAATACGTC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGCGTTACCGGTTTGT	XhoI
597	Forward	CGCGGATCCCATATG-CTGCTTCATGTCAGC	BamHI- NdeI
	Reverse	GCCCAAGCTT-ACGTATCCAGCTCGAAG	HindIII
601	Forward	CGCGGATCCCATATG-ATATGTTCCCAACCGGCAAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAAACAATCCTCAGGCAC	XhoI
602	Forward	CGCGGATCCGCTAGC-TTGCTCCATCAATGC	BamHI- NheI
	Reverse	CCCGCTCGAG-ATGCAGCTGCTAAAAGCG	XhoI
603	Forward	AAAAAAGAATTC-CTGTCCTCGCGTAGGCGGGGACGGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAGATGCCGGCAAGTTCGGC	Pst I
604	Forward	CGCGGATCCGCTAGC-CCCGAAGCGCACTT	BamHI- NheI
	Reverse	CCCGCTCGAG-GACGGCATCTGCACGG	XhoI
606a	Forward	AAAAAAGAATTC-CGCGAATACCGCGCCGATGCGGGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAGCGATTTGAGGCGGGCGATACG	Pst I
607	Forward	AAAAAAGAATTC-ATGCTGCTCGACCTCAACCGCTTTTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCCTTATGCGATCTGAC	Pst I
608	Forward	AAAAAAGAATTC-ATGTCCGCCCTCCTCCCCATCATCAACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTCTATCCAAATGTCGCGTTC	Pst I
609	Forward	CGCGGATCCCATATG-GTTGTGGATAGACTCG	BamHI- NdeI

	Reverse CCCGCTCGAG-CTGGATTATGATGTCTGTC	XhoI
610	Forward CGCGGATCCCATATG-ATTGGAGGGCTTATGCA	BamHI- NdeI
	Reverse CCCGCTCGAG-ACGCTTCAACATCTTTGCC	XhoI
611	Forward CGCGGATCCCATATG-CCGTCTCAAAACGGG	BamHI- NdeI
	Reverse CCCGCTCGAG-AACGACTTTGAACGCGCAA	XhoI
613	Forward CGCGGATCCCATATG-TCGCGTTCGAGCCG3	BamHI- NdeI
	Reverse CCCGCTCGAG-AGCCTGTAAAATAAGCGGC	XhoI
614	Forward CGCGGATCCCATATG-TCCGTCGTGAGCGGC	BamHI- NdeI
	Reverse CCCGCTCGAG-CCATACTGCGGCGTTC	XhoI
616	Forward AAAAAAGAATTC-ATGTCAAACACAATCAAAATGGTTGTCGG	Eco RI
	Reverse AAAAAATCTAGA-TTAGTCCGGGCGGCAGGCAGCTCG	Xba I
619a	Forward AAAAAAGAATTC-GGGCTTCTCGCCGCCTCGCTTGC	Eco RI
	Reverse AAAAAACTGCAG-TCATTTTTTTGTGTTTTTAAAACGAGATA	Pst I
622	Forward CGCGGATCCCATATG-GCCGCCCTGCCTAAAG	BamHI- NdeI
	Reverse CCCGCTCGAG-TTTGTCCAAATGATAAATCTG	XhoI
624	Forward CGCGGATCCCATATG-TCCCCGCGCTTTTACCG	BamHI- NdeI
	Reverse CCCGCTCGAG-AGATTCGGGCCTGCGC	XhoI
625	Forward CGCGGATCCCATATG-TTTGCAACCAGGAAAATG	BamHI- NdeI
	Reverse CCCGCTCGAG-CGGCAAATTACCGCCTT	XhoI
627a	Forward AAAAAAGAATTC-AAAGCAGGCGAGGCAGGCGCGCTGGG	Eco RI
	Reverse AAAAAACTGCAG- TTACGAATGAAACAGGGTACCCGTCATCAAGGC	Pst I
628	Forward AAAAAAGGTACC-GCCTTACAAACATGGATTTTGCGTTC	Kpn I
	Reverse AAAAAACTGCAG-CTACGCACCTGAAGCGCTGGCAA	Pst I
629a	Forward AAAAAAGAATTC-GCCACCTTTATCGCGTATGAAAACGA	Eco RI
	Reverse AAAAAACTGCAG-TTACAACACCGCCGTCCGGTTCAAACC	Pst I
630a	Forward AAAAAAGAATTC-GCGGCTTTGGGTATTTCTTTTCGG	Eco RI
	Reverse AAAAAACTGCAG-TTAGGAGACTTCGCCAATGGAGCCGGG	Pst I
635	Forward AAAAAAGAATTC- ATGACCCAGCGACGGGTTCGGCAAGCAAAACCG	Eco RI
	Reverse AAAAAACTGCAG-TTAATCCACTATAATCCTGTTGCT	Pst I
638	Forward AAAAAAGAATTC-ATGATTGGCGAAAAGTTTATCGTAGTTGG	Eco RI
	Reverse AAAAAACTGCAG-TCACGAACCGATTATGCTGATCGG	Pst I
639	Forward CGCGGATCCCATATG-ATGCTTTATTTTGTTCG	BamHI- NdeI
	Reverse CCCGCTCGAG-ATCGCGGCTGCCGAC	XhoI
642	Forward CGCGGATCCCATATG-CGGTATCCGCCGCAAT	BamHI- NdeI
	Reverse CCCGCTCGAG-AGGATTGCGGGGCATTA	XhoI

643	Forward	CGCGGATCCCATATG-GCTTCGCCGTCGGCAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACCGAAAAACAGACCGC	XhoI
644	Forward	AAAAAAGAATTC- ATGCCGTCTGAAAGGTCGGCGGATTGTTGCCC	Eco RI
	Reverse	AAAAAATCTAGA-CTACCCGCAATATCGGCAGTCCAATAT	Pst I
645	Forward	AAAAAAGAATTC-GTGGAACAGAGCAACACGTAAATCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGAGGAAACCGAAGACCAGGCCGC	Pst I
647	Forward	AAAAAAGAATTC-ATGCAAAGGCTCGCCGCAGACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGATTATCAGGGATATCCGGTAGAA	Pst I
648	Forward	AAAAAAGAATTC- ATGAACAGGCGCGACGCGCGGATCGAACG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCTGTGTGCTGATTGAATGCGAC	Pst I
649	Forward	AAAAAAGAATTC-GGTACGTCAGAACCCGCCACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGCGGAAACTGCCGCCGTC	Pst I
650	Forward	AAAAAAGAATTC-ATGTCCAAACTCAAACCATCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCATGGCGGTCTGTTTT	Pst I
652	Forward	AAAAAAGGTACC- GCTGCCGAAGACTCAGGCCTGCCGCTTTACCG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTTGCCAGTTGGTAGAATGCGGC	Pst I
653	Forward	AAAAAAGAATTC-GCGGCTTTGCCGGTAATTTTCATCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTATGCCGGTCTGGTTGCCGGCGGCGA	Pst I
656a	Forward	AAAAAAGAATTC-CGGCCGACGTCGTTGCGTCCTAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-CTACGATTTCGGCGATTTCACATCGT	Pst I
657	Forward	AAAAAAGAATTC-GCAGAATTTGCCGACCGCCATTTGTGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATAGGGACTGATGCAGTTTTTTTGC	Pst I
658	Forward	CGCGGATCCCATATG-GTGTCCGGAATTGTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCAGAATGTTTACCGTT	XhoI
661	Forward	AAAAAAGAATTC- ATGCACATCGGCGGCTATTTTATCGACAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TCACGACGTGTCTGTTCCGCCGTCGGGC	Pst I
663	Forward	CGCGGATCCCATATG-TGTATCGAGATGAAATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTAAAAATCGGGGCTGC	XhoI
664	Forward	CGCGGATCCCATATG-GCGGCTGGCGCGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATCGAGTTTTACACCAC	XhoI
665	Forward	AAAAAAGAATTC-ATGAAATGGGACGAAACGCGCTTCGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAAAATTTGCCGACGATTTC	Pst I
666	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
667	Forward	AAAAAAGAATTC- CCGCATCCGTTTGATTTCATTTCGTATTCGTCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAATGACACAATAGGCGCAAGTC	Pst I

669	Forward	AAAAAAGAATTC-ATGCGCCGCATCATTA	AAAAAACACCAGCC	Eco RI
	Reverse	AAAAAACTGCAG-TTACAGTATCCGTTT	GATGTCGGC	Pst I
670a	Forward	AAAAAAGAATTC-AAAAACGCTTCGGGCGTTTC	GTCTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGAGCTTTT	GGAACGCGTCGGACTGGC	Pst I
671	Forward	CGCGGATCCCATATG-ACCAGCAGGGTAAC		BamHI-
	Reverse	CCCGCTCGAG-AGCAACTATAAAAAACGCAAG		NdeI
672	Forward	CGCGGATCCCATATG-AGGAAAATCCGCACC		XhoI
	Reverse	CCCGCTCGAG-ACGGGATAGGCGGTTG		BamHI-
673	Forward	AAAAAAGAATTC-ATGGATATTGAAACCTTCCTT	GCAGG	NdeI
	Reverse	AAAAAACTGCAG-CTACAAACCCAGCTCGCGCAGGAA		XhoI
674	Forward	AAAAAAGAATTC-ATGAAAACAGCCCGCCGCGTTCCCG		Eco RI
	Reverse	AAAAAACTGCAG-TCAACGGCGTTTGGGCTCGTCGGG		Pst I
675	Forward	CGCGGATCCCATATG-AACACCATCGCCCC		Eco RI
	Reverse	CCCGCTCGAG-TTCTTCGTCTTCAAACCTGT		Pst I
677a	Forward	AAAAAAGAATTC-AGACGGCATTC	CCCGATCAGTCGATTTTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGTATGCGCGAAATCGACCGCCGC		Pst I
680	Forward	CGCGGATCCGCTAGC-ACGAAGGGCAGTTCGG		BamHI-
	Reverse	CCCGCTCGAG-CATCAAAAACCTGCCGC		NheI
681	Forward	AAAAAAGAATTC-ATGACGACGCCGATGGCAATCAGTGC		XhoI
	Reverse	AAAAAACTGCAG-TTACCGTCTTCCGCAAAAAACAGC		Eco RI
683	Forward	CGCGGATCCCATATG-TGCAGCACACCGGACAA		Pst I
	Reverse	CCCGCTCGAG-GAGTTTTTTTCCGCATACG		BamHI-
684	Forward	CGCGGATCCCATATG-TGCGGTACTGTGCAAAG		NdeI
	Reverse	CCCGCTCGAG-CTCGACCATCTGTTGCG		XhoI
685	Forward	CGCGGATCCCATATG-TGTTTGCTTAATAATAAACATT		BamHI-
	Reverse	CCCGCTCGAG-CTTTTTCCCCGCCGCA		NdeI
686	Forward	CGCGGATCCCATATG-TGCGGCGGTTCGGAAG		XhoI
	Reverse	CCCGCTCGAG-CATTCCGATTCTGATGAAG		BamHI-
687	Forward	CGCGGATCCCATATG-TGCGACAGCAAAGTCCA		NdeI
	Reverse	CCCGCTCGAG-CTGCGCGGCTTTTTTGTT		XhoI
690	Forward	CGCGGATCCCATATG-TGTTCTCCGAGCAAAGAC		BamHI-
	Reverse	CCCGCTCGAG-TATTCGCCCCGTGTTTGG		NdeI
691	Forward	CGCGGATCCCATATG-GCCACGGCTTATATCCC		XhoI
	Reverse	CCCGCTCGAG-TTTGAGGCAGGAAGAAAG		BamHI-
				NdeI
				XhoI

694	Forward	CGCGGATCCCATATG-TTGGTTTCCGCATCCGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCTGCGTCGGTGCGGT	XhoI
695	Forward	CGCGGATCCCATATG-TTGCCTCAAACCTCGTCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
696	Forward	CGCGGATCCCATATG-TTGGGTTGCCGGCAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGATTGCCGCAATGATG	XhoI
700a	Forward	AAAAAAGAATTC-GCATCGACAGACGGTGTGTCGTGGAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCTACCGGCACGACTTCCAAACC	Pst I
701	Forward	CGCGGATCCCATATG-AAGACTTGTTTGGATACTTC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCCGACAACAGCCTC	XhoI
702	Forward	AAAAAAGAATTC-ATGCCGTGTTCCAAAGCCAGTTGGATTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACCCCATTCACCCGGAGAACCGA	Pst I
703	Forward	CGCGGATCCGCTAGC-CAAACGCTGGCAACCG	BamHI- NheI
	Reverse	CCCGCTCGAG-TTTTGCAGGTTTGATGTTTG	XhoI
704a	Forward	AAAAAAGAATTC-GCTTCTACCGGTACGCTGGCGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTTTTGCCGGATAATATGGCGGGTGCG	Pst I
707	Forward	CGCGGATCCGCTAGC-GAAATTATTAACGATGCAGA	BamHI- NheI
	Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGA	XhoI
708	Forward	CGCGGATCCGCTAGC-CCTTTTAAGCCATCCAAAA	BamHI- NheI
	Reverse	CCCGCTCGAG-TTGACCGGTGAGGACG	XhoI
710	Forward	CGCGGATCCCATATG-GAAACCCACGAAAAAATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACGGTTTCGGTCAG	XhoI
714	Forward	CGCGGATCCCATATG-AGCTATCAAGACATCTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGTAGGTAAATCGGAT	XhoI
716	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
718	Forward	CGCGGATCCCATATG-GAGCCGATAATGGCAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCGCGGGCATGGTCTTGTC	XhoI
720	Forward	CGCGGATCCCATATG-AGCGGATGGCATAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTTGCATAGCTGTTGACCA	XhoI
723	Forward	CGCGGATCCCATATG-CGACCCAAGCCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AATGCGAATCCGCCGCC	XhoI

725	Forward	CGCGGATCCCATATG-GTGCGCACGGTTAAA	BamHI-
	Reverse	CCCGCTCGAG-TTGCTTATCCTTAAGGGTTA	NdeI
726	Forward	CGCGGATCCCATATG-ACCATCTATTTCAAAAAC	XhoI
	Reverse	CCCGCTCGAG-GCCGATGTTTAGCGTCC	BamHI-
728	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	NdeI
	Reverse	CCCGCTCGAG-GTGAGAAAGGTCGCGC	XhoI
729	Forward	CGCGGATCCCATATG-TGCACCATGATTCCCCA	BamHI-
	Reverse	GCCCAAGCTT-TTGTCGGTTTGGGTATC	NdeI
731	Forward	CGCGGATCCGCTAGC-GCCGTGCCGGAGG	HindIII
	Reverse	CCCGCTCGAG-ACGGGCGCGGCAG	BamHI-
732	Forward	CCGGAATTCTACATATG-TCGAAACCTGTTTTTAAGAA	NheI
	Reverse	CCCGCTCGAG-CTTCTTATCTTTTTTATCTTTC	XhoI
733	Forward	CGCGGATCCCATATG-GCCTGCGGCGGCAA	EcoRI-
	Reverse	CCCGCTCGAG-TCGCTTGCCTCCTTTAC	NdeI
734	Forward	CGCGGATCCCATATG-GCCGATACTTACGGCTAT	XhoI
	Reverse	CCCGCTCGAG-TTTGAGATTTTGAATCAAAGAG	BamHI-
735	Forward	CGCGGATCCCATATG-AAGCAGCAGGCGGTCA	NdeI
	Reverse	CCCGCTCGAG-ATTTCCGTAGCCGAGGG	XhoI
737	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	BamHI-
	Reverse	CCCGCTCGAG-GTCGTCGCGGCGGGA	NdeI
739	Forward	CGCGGATCCCATATG-GCAAAAAAACC GAACA	XhoI
	Reverse	CCCGCTCGAG-GAAGAGTTTGTCGAGAATT	BamHI-
740	Forward	CGCGGATCCCATATG-GCCAATCCGCCCGAAG	NdeI
	Reverse	CCCGCTCGAG-AAACGCGCCAAAATAGTG	XhoI
741	Forward	CGCGGATCCCATATG-TGCAGCAGCGGAGGG	BamHI-
	Reverse	CCCGCTCGAG-TTGCTTGGCGGCAAGGC	NdeI
743	Forward	CGCGGATCCCATATG-GACGGTGTTGTGCCTGTT	XhoI
	Reverse	CCCGCTCGAG-CTTACGGATCAAATTGACG	BamHI-
745	Forward	CGCGGATCCCATATG-TTTTGGCAACTGACCG	NdeI
	Reverse	CCCGCTCGAG-CAAATCAGATGCCTTTAGG	XhoI
746	Forward	CGCGGATCCCATATG-TCCGAAAACAAACAAAAC	BamHI-

	Reverse	CCCGCTCGAG-TTCATTCGTTACCTGACC	NdeI
747	Forward	CCGGAATTCTAGCTAGC-CTGACCCCTTGGG	XhoI
	Reverse	GCCCAAGCTT-TTTTGATTTTAATTGACTATAGAAC	EcoRI-
749	Forward	CGCGGATCCCATATG-TGCCAGCCGCCG	NheI
	Reverse	CCCGCTCGAG-TTTCAAGCCGAGTATGC	HindIII
750	Forward	CGCGGATCCCATATG-TGTTCGCCCGAACCTG	BamHI-
	Reverse	CCCGCTCGAG-CTTTTTCCCCGCCGCAA	NdeI
758	Forward	CGCGGATCCCATATG-AACAATCTGACCGTGTT	XhoI
	Reverse	CCCGCTCGAG-TGGCTCAATCCTTTCTGC	BamHI-
759	Forward	CGCGGATCCGCTAGC-CGCTTCACACACACCAC	NdeI
	Reverse	CCCGCTCGAG-CCAGTTGTAGCCTATTTTG	XhoI
763	Forward	CGCGGATCCCATATG-CTGCCTGAAGCATGGCG	BamHI-
	Reverse	CCCGCTCGAG-TTCCGCAAATACCGTTTCC	NdeI
764	Forward	CGCGGATCCCATATG-TTTTCTCCGCCCTGA	XhoI
	Reverse	CCCGCTCGAG-TCGCTCCCTAAAGCTTTC	BamHI-
765	Forward	CGCGGATCCCATATG-TTAAGATGCCGTCCG	NdeI
	Reverse	CCCGCTCGAG-ACGCCGACGTTTTTTATTAA	XhoI
767	Forward	CGCGGATCCCATATG-CTGACGGAAGGGGAAG	BamHI-
	Reverse	CCCGCTCGAG-TTTCTGTACAGCAGGGG	NdeI
768	Forward	CGCGGATCCCATATG-GCCCCGCAAAAACCCG	XhoI
	Reverse	CCCGCTCGAG-TTTCATCCCTTTTTTGAGC	BamHI-
770	Forward	CGCGGATCCCATATG-TGCGGCAGCGGCGAA	NdeI
	Reverse	CCCGCTCGAG-GCGTTTGTCGAGATTTTC	XhoI
771	Forward	CGCGGATCCCATATG-TCCGTATATCGCACCTTC	BamHI-
	Reverse	CCCGCTCGAG-CGGTTCTTTAGGTTTGAG	NdeI
772	Forward	CGCGGATCCCATATG-TTGCGGCGTTGGTGG	XhoI
	Reverse	CCCGCTCGAG-CAATGCCGACATCAAACG	BamHI-
774	Forward	CGCGGATCCCATATG-TCCGTTTCACCCGTTCC	NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
790	Forward	CGCGGATCCCATATG-GCAAGAAGGTCAAAAAC	BamHI-
			NdeI

	Reverse CCCGCTCGAG-GGCGTTGTTCCGATTTCG	XhoI
900	Forward CGCGGATCCCATATG-CCGTCTGAAATGCCG	BamHI- NdeI
	Reverse CCCGCTCGAG-ATATGGAAAAGTCTGTTGTC	XhoI
901	Forward CGCGGATCCCATATG-CCCGATTTTTCGATG	BamHI- NdeI
	Reverse CCCGCTCGAG-AAAATGGAACAATACCAGG	XhoI
902	Forward. CCGGAATTCTACATATG-TTGCACCTTCAAAGGATAATC	EcoRI- NdeI
2		
	Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC	XhoI
903	Forward CCGGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT	EcoRI- NheI
	Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA	XhoI
904	Forward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC	Kpn I
	Reverse AAAGTGCAG-TTAATATCGATAGGTTATATG	Pst I
904a	Forward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA	Eco RI
	Reverse AAAGTGCAG-TTAATATCGATAGGTTATATG	Pst I
905	Forward CGCGGATCCCATATG-AACAAAATATACCGCATC	BamHI- NdeI
	Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT	XhoI
907	Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG	BamHI- NdeI
	Reverse CCCGCTCGAG-ACGCCACTGCCAGCG	XhoI
908	Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC	Eco RI
	Reverse AAAGTGCAG-TTAATATGGTTTTGTCTGTTTCG	Pst I
909	Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT	BamHI- NdeI
	Reverse CCCGCTCGAG-TCGGTTTTGAACTTTGGTTTT	XhoI
910	Forward AAAGAATTC-GCATTGCGCGGCGACTCTGCCGAGCG	Eco RI
	Reverse AAAGTGCAG-TCAGCGATCGAGCTGCTCTTT	Pst I
911	Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC	Eco RI
	Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC	Pst I
912	Forward AAAAAAGAATTC- CAAATCCGTCAAAACGCCACTCAAGTATTGAG	Eco RI
	Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC	Pst I
913	Forward CGCGGATCCCATATG-GAAACCCGCCCGC	BamHI- NdeI
	Reverse CCCGCTCGAG-AGGTTGTGTTCCAGGTTG	XhoI
915	Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA	BamHI- NdeI
	Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG	XhoI
914	Forward AAAGAATTC-GACAGAATCGGCGATTGGAAGCACG	Eco RI
	Reverse AAAGTGCAG-CTATATGCGCGGACGCTCAACGG	Pst I
916	Forward CGCGGATCCCATATG-GCAATGATGGCGGCTG	BamHI- NdeI
	Reverse CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI



917	Forward	AAAAAAGAATTC-CCTGCCGAAAAACCGGCACCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCCCCGCCTTCACATCCTG	Pst I
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC	BamHI-
	Reverse	CCCGCTCGAG-CGGGCGGTATTCTGGG	NdeI
920	Forward	CGCGGATCCCATATG-CACCGCGTCTGGGTC	XhoI
	Reverse	CCCGCTCGAG-ATGGTGCGAATGACCGA	BamHI-
921	Forward	AAAAAAGAATTC-TTGACGGAAATCCCCGTGAATCC	NdeI
	Reverse	AAAAAACTGCAG-TCATTTCAAGGGCTGCATCTTCAT	XhoI
922	Forward	CGCGGATCCGCTAGC-TGTACGGCGATGGAGGC	Eco RI
	Reverse	CCCGCTCGAG-CAATCCCAGGCGCC	Pst I
923	Forward	CGCGGATCCCATATG-TGTTACGCAATATTGTCCC	BamHI-
	Reverse	CCCGCTCGAG-GGACAAGGCGACGAAG	NheI
925	Forward	CGCGGATCCCATATG-AAACAAATGCTTTTAGCCG	XhoI
	Reverse	CCCGCTCGAG-GCCGTTGCATTTGATTTC	BamHI-
926	Forward	CGCGGATCCCATATG-TGCGCGCAATTACCTC	NdeI
	Reverse	CCCGCTCGAG-TCTCGTGCGCGCCG	XhoI
927	Forward	CGCGGATCCCATATG-TGCAGCCCCGCAGC	BamHI-
	Reverse	CCCGCTCGAG-GTTTTTTGCTGACGTAGT	NdeI
929a	Forward	AAAAAAGAATTC-CGCGGTTTGCTCAAAACAGGGCTGGG	XhoI
	Reverse	AAAAAATCTAGA-TTAAGAAAGACGGAACTACTGCC	Eco RI
931	Forward	AAAAAAGAATTC-GCAACCCATGTTTTGATGGAAAC	Xba I
	Reverse	AAAAAACTGCAG-TTACTGCCCGACAACAACGCGACG	Eco RI
935	Forward	AAAAAAGAATTC-GCGGATGCGCCCGCGATTTTGGATGACAAGGC	Pst I
	Reverse	AAAAAACTGCAG-TCAAACCGCCAATCCGCCGACAC	Eco RI
936	Forward	CGCGGATCCCATATG-GCCGCCGTCGGCGC	Pst I
	Reverse	CCCGCTCGAG-GCGTTGGACGTAGTTTTG	BamHI-
937	Forward	AAAAAAGAATTC-CCGGTTTACATTCAAACCGGCGCAAC	NdeI
	Reverse	AAAAAACTGCAG-TTAAAATGTATGCTGTACGCCAAA	XhoI
939a	Forward	AAAAAAGAATTC-GGTTTCGGCAGCTGTGATGAAACC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGCAAACCTTGGATAAAGTTGGC	Pst I
950	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	Eco RI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	BamHI-
953	Forward	CGCGGATCCCATATG-GCCACCTACAAAGTGGAC	NdeI
	Reverse	CCCGCTCGAG-TTGTTTGGCTGCCTCGAT	XhoI
957	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI-

	Reverse CCCGCTCGAG-GTGAGAAAGGTCGCGC	NdeI
958	Forward CGCGGATCCCATATG-GCCGATGCCGTTGCG	XhoI
		BamHI-
	Reverse GCCCAAGCTT-GGGTCGTTTGTGCGTC	NdeI
959	Forward CGCGGATCCCATATG-CACCACGACGGACACG	HindIII
		BamHI-
	Reverse CCCGCTCGAG-GTCGTCGCGGCGGGA	NdeI
961	Forward CGCGGATCCCATATG-GCCACAAGCGACGACG	XhoI
		BamHI-
	Reverse CCCGCTCGAG-CCACTCGTAATTGACGC	NdeI
972	Forward AAAAAAGAATTC-	XhoI
	TTGACTAACAGGGGGGGGAGCGAAATTAAAAAC	Eco RI
	Reverse AAAAAATCTAGA-TTAAAAATAATCATAATCTACATTTTG	
973	Forward AAAAAAGAATTC-ATGGACGGCGCACAACCGAAAAC	Xba I
	Reverse AAAAACTGCAG-TTACTTCACGCGGGTTCGCCATCAGCGT	Eco RI
982	Forward CGCGGATCCCATATG-GCAGCAAAGACGTAC	Pst I
		BamHI-
	Reverse CCCGCTCGAG-CATCATGCCGCCCATCC	NdeI
983	Forward CGCGGATCCCATATG-TTAGCTGTTGCAACAACAC	XhoI
		BamHI-
	Reverse CCCGCTCGAG-GAACCGGTAGCCTACG	NdeI
987	Forward CGCGGATCCCATATG-CCCCACTGGAAGAAC	XhoI
		BamHI-
	Reverse CCCGCTCGAG-TAATAAACCTTCTATGGGC	NdeI
988	Forward CGCGGATCCCATATG-TCTTTAAATTTACGGGAAAAAG	XhoI
		BamHI-
	Reverse GCCCAAGCTT-TGATTTGCCTTTCCGTTTT	NdeI
989	Forward CCGGAATTCTACATATG-GTCCACGCATCCGGCTA	HindIII
		EcoRI-
	Reverse CCCGCTCGAG-TTTGAATTTGTAGGTGTATTGC	NdeI
990	Forward. CGCGGATCCGCTAGC-TTCAGAGCTCAGCTT	XhoI
	2	BamHI-
	Reverse CCCGCTCGAG-AAACAGCCATTTGAGCGA	NheI
992	Forward CGCGGATCCCATATG-GACGCGCCCGCCCG	XhoI
		BamHI-
	Reverse CCCGCTCGAG-CCAAATGCCCAACCATTC	NdeI
993	Forward CGCGGATCCCATATG-GCAATGCTGATTGAAATCA	XhoI
		BamHI-
	Reverse CCCGCTCGAG-GAACACATCGCGCCCG	NdeI
996	Forward CGCGGATCCCATATG-TGCGGCAGAAAATCCGC	XhoI
		BamHI-
	Reverse CCCGCTCGAG-TCTAAACCCCTGTTTTCTC	NdeI
997	Forward CCGGAATTCTAGCTAGC-CGGCACGCCGACGTT	XhoI
		EcoRI-
	Reverse CCCGCTCGAG-GACGGCATCGCTCAGG	NheI
		XhoI

Underlined sequences indicate restriction recognition sites.

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 1>:

```
g001.seq
  1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG GTGTCGGCGA ACGAGGTGTC
 51  CGGCAGGGCT TCGCGCCGGA TGGTGCTGGT CATCTGCCAG ACGCTGCCGA
101  AACGCGATAC TTAAACGGC TCGGGTACGC ATACTTTACC GGTTTGGGCG
151  ATTTTGCCGA GGTCGTTGCG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201  GCGGTTTTTC GGGTCGGTTT GTAACTCGGC GCGCGGCGCT TCGTCTTGTC
251  CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCTGGT GCTGATGGTG
301  CCGTCTGAAG CGATGTTGAG GAAGAGTTCG GCGGAGAAAC ACAGCGTCCA
351  CGCGGATTGC CCGGCTTCAT CGGGCAGGTG GGACAATACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 2; ORF 001.ng>:

```
g001.pep
  1  MLPQGKAARR VSANEVSGRA CARMVLVICQ TLPKRDTLNG SGTHTLPVWA
 51  ILPRSLRSKS TIITFSARFF GSVCSAARR SSCPSPKIGA VPFIGSVLMV
101  PSEAMLRKSS GEKHSVHAD C PASSGRWDNT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3>:

```
m001.seq
  1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTC
 51  CGGcAssCTT ss.GCTTGGGA yGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101  AACGCGATAC TTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
151  ATTTTGCCGA GATCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201  GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GCGCGGCGCT TCGTCTTGTC
251  CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCTGGT GCTGATGGTG
301  CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GCGGAGAAAC ACAGCGTCCA
351  CGCGGATTGC CCTCCGCAT CGGGCAGGTG GGACAAGACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 4; ORF 001>:

```
m001.pep
  1  MLPQGKAARR MSANEVCXL XAWXVLVICQ TLPKRDTLNG SGTHTVPVWA
 51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101  PSEPILRKSS GEKHSVHAD C PSASGRWDKT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 5>:

```
a001.seq
  1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTC
 51  CGGCAAGGCT TGGGCTTGGGA TGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101  AACGCGATAC TTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
151  ATTTTGCCGA GGTCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201  GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GCGCGGCGCT TCGTCTTGTC
251  CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCTGGT GCTGATGGTG
301  CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GCGGAGAAAC ACAGCGTCCA
351  CGCGGATTGC CCTTGTGCAT CGGGCAGGTG GGACAAAACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 6; ORF 001.a>:

```
a001.pep
  1  MLPQGKAARR MSANEVCGKA WAWMVLVICQ TLPKRDTLNG SGTHTVPVWA
 51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101  PSEPILRKSS GEKHSVHAD C PCASGRWDKT A*
```

m001/a001 96.2% identity over a 131 aa overlap

```

              10      20      30      40      50      60
m001.pep  MLPQGKAARRMSANEVCGXLXAWXVLVICQLPKRDTLNGSGTHTVPVWAILPRSLRSKS
          |||
a001.pep  MLPQGKAARRMSANEVCGKAWAWMVLVICQLPKRDTLNGSGTHTVPVWAILPRSLRSKS
```

	10	20	30	40	50	60
	70	80	90	100	110	120
m001.pep	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
a001.pep	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
	70	80	90	100	110	120
	130					
m001.pep	PSASGRWDKTAX					
a001.pep	PCASGRWDKTAX					
	130					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 001 shows 89.3% identity over a 131 aa overlap with a predicted ORF (ORF 001.ng) from *N. gonorrhoeae*:

m001/g001

	10	20	30	40	50	60
m001.pep	MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGHTHPVWAILPRSLRSKS					
g001	MLPQGKAARRVSANEVSGRACARMVLVICQTLPKRDTLNGSGHTHPVWAILPRSLRSKS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m001.pep	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
g001	TIIITFSARFFGSVCNSAARRSSCPSPKIGAVPFIGSVLMVPSEAMLRKSSGEKHSVHADC					
	70	80	90	100	110	120
	130					
m001.pep	PSASGRWDKTAX					
g001	PASSGRWDNTAX					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 7>:

g003.seq

```

1   ATGGTCGTAT TCGTGGCTGA AGGCGTATTC GGTCGCGCTG TTTTGGGTCA
51  CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACTCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCAGTG CGGCTTTGGT
151 TTTGCCCGGC AGCGGTTCTG CGGCTTTGCG GATGTCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAGG TAGTCCTGAT GGTATTCCTC GGCGTCGTAG
251 AAGTTTTTCA GCGGTTCTGT TTCAACAACG AGGGGCAGTT GGTATTTTGT
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
351 TGAGTACAC GCCGCTGCGG TATTGCGTGC CGGTGTCGTT ACCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAA TAATATTGCA GGATGTCGTC
451 CAGgCTGagt TTGTCGGCAT CGTaggtcac tTTGACGGTC TCGGCATGAC
501 CCGTATGGCG GTaggacact tctTCgtanc TcGGGtTTTC CGTGttGCCG
551 TTGGCgttac cGGATACCGC gteaACCACG CCGTcgatgc gttggaAATa
601 ggCTTCCAAG ccccaaaagc agccgcgcgc gaagtaaata gtgcccgtgt
651 tcatgattGC TGa

```

This corresponds to the amino acid sequence <SEQ ID 8; ORF 003.ng>:

g003.pep

```

1   MVVFVAEGVF GRAVLGHLVL LFGQGAFFG VTRFFIRCRV EAFALRCGFG
51  FARQRFVGF DVDVAVAVGV FNQVVLMLVL GVVEVFQRFV FNNEGQLVFL
101 LLAFEGGGDD GFFGGVGVVH AAVLRAGVV TLFVEAGRIN DAEIILQDVV

```

151 QAEFVGIVGH FDGLGMTRMA VGHFFVRVFR VAVGVGTGYRV NHAVDAL EIG  
201 FQAPKAAAGE VNGARVHDC

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 9>:

m003.seq  
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTGCGCTG TTTTGGGTAA  
51 CTTGStATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCAC TCGGT  
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGGG CCGTCTTGGT  
151 TTTGCCCGGC AGCGGTTCGT CAGCKTTGCG GATGTCGATG TGGCAGTAGC  
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG  
251 AAGTTTTtCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTGT  
301 CTGCTCGCGT TTGAGGGCGk CGGCGATGAC GGCTTTTTTCG kCGGGGTCGG  
351 TGTAGTACAC GCCGCTGCGG TATTGCGTAC CCGTGTCGTT GCCCTGTTTG  
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC  
451 TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGCG  
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTTGCCG  
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAATA  
601 GGCTTCCAAG CCCCAGAAGC AGCg.CCGGC GAGGTAAATG GTGCGCGTGT  
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 10; ORF 003>:

m003.pep Length: 221  
1 MVVFVAEGIF GRAVLGNLXL LFGQGAFFFG VTRFFIRCRV EAFALRGGLG  
51 FARQRFVSXA DVDVAVAVGV FNQVVL MVFL GIVEVFQRLV FNNEGQLVFL  
101 LLAFEGXGDD GFFXGVGVVH AAVLRTGVV ALFVEAGRIN DAEEILQDVV  
151 \*AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDAL E I  
201 GFQAPEAAXG EVNGARVHDF \*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 11>:

a003.seq  
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTGCGCTG TTTTGGGTAA  
51 CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCAC TCGGT  
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CCGTCTTGGT  
151 TTTGCCCGGC AGCGGTTCGT CGGCTTTGCG GATATCGATG TGGCAGTAGC  
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG  
251 AAGTTTTTCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTGT  
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTTCG GCGGGGTCGG  
351 TGTAGTACAC GCCGCTGCGG TATTGCGTAC CCGTGTCGTT GCCCTGTTTG  
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC  
451 TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGCG  
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTTGCCG  
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAATA  
601 GGCTTCCAAG CCCCAGAAGC AGCCGCCGGC GAGGTAGATG GTGCGCGTGT  
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 12; ORF 003.a>:

a003.pep  
1 MVVFVAEGIF GRAVLGNLVL LFGQGAFFFG VTRFFIRCRV EAFALRCGLG  
51 FARQRFVGFA DIDVAVAVGV FNQVVL MVFL GIVEVFQRLV FNNEGQLVFL  
101 LLAFEGGGDD GFFGGGVGVVH AAVLRTGVV ALFVEAGRIN DAEEILQDVV  
151 \*AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDAL E I  
201 GFQAPEAAAAG EVDGARVHDF \*

m003/a003 95.9% identity over a 220 aa overlap

	10	20	30	40	50	60
m003.pep	MVVFVAEGIFGRAVLGNLXL LFGQGAFFFGVTRFFIRCRVEAFALRGGLGFARQRFVSXA					
a003	MVVFVAEGIFGRAVLGNLVL LFGQGAFFFGVTRFFIRCRVEAFALRCGLGFARQRFVGFA					
	10	20	30	40	50	60

m003.pep	70	80	90	100	110	120
	DVDVAVAVGVFNQVVL	MLVFLGIVEVFQRLV	FNNEGQLVFLLLAFEG	XGDDGFFXGVGVVH		
a003	70	80	90	100	110	120
	DIDVAVAVGVFNQVVL	MLVFLGIVEVFQRLV	FNNEGQLVFLLLAFEG	GGDDGFFXGVGVVH		
m003.pep	130	140	150	160	170	180
	AAAVLRTGVVALFVEAG	RINDAEIILQDVVXAE	FVGIVGHFDGFGVARM	AVGHVFIARIF		
a003	130	140	150	160	170	180
	AAAVLRTGVVALFVEAG	RINDAEIILQDVVXAE	FVGIVGHFDGFGVARM	AVGHVFIARIF		
m003.pep	190	200	210	220		
	RVAVGVAGYRVNHA	VDALIGFQAPEAA	XGEVNGARVHDFX			
a003	190	200	210	220		
	RVAVGVAGYRVNHA	VDALIGFQAPEAA	AGEVDGARVHDFX			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 003 shows 88.6% identity over a 219 aa overlap with a predicted ORF (ORF 003.ng) from *N. gonorrhoeae*:

m003/g003

m003.pep	10	20	30	40	50	60
	MVVFVAEGIFGRAVL	GNLXLLFGQGA	FEFGVTRFFI	RCRVEAFALR	GGLGFARQ	RFVSXA
g003	10	20	30	40	50	60
	MVVFVAEGVFGRAVL	GHLVLLFGQGA	FEFGVTRFFI	RCRVEAFALR	CGFGFARQ	RFVGF
m003.pep	70	80	90	100	110	120
	DVDVAVAVGVFNQVVL	MLVFLGIVEVFQRLV	FNNEGQLVFLLLAFEG	XGDDGFFXGVGVVH		
g003	70	80	90	100	110	120
	DVDVAVAVGVFNQVVL	MLVFLGIVEVFQRLV	FNNEGQLVFLLLAFEG	GGDDGFFXGVGVVH		
m003.pep	130	140	150	160	170	180
	AAAVLRTGVVALFVEAG	RINDAEIILQDVVXAE	FVGIVGHFDGFGVARM	AVGHVFIARIF		
g003	130	140	150	160	170	180
	AAAVLRAGVVTLFVEAG	RINDAEIILQDVVQAE	FVGIVGHFDGLGMTR	MAVGHFFV-RVF		
m003.pep	190	200	210	220		
	RVAVGVAGYRVNHA	VDALIGFQAPEAA	XGEVNGARVHDFX			
g003	190	200	210			
	RVAVGVAGYRVNHA	VDALIGFQAPKAA	AGEVNGARVHDC			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 13>:

g004.seq

1	ATGgtagAAC	GGCATATCCA	GCATTGCGG	AACGGTCATC	TTCATTGAT
51	GCGCCCATGC	CAACAagtga	gccAAAtgtT	CGGCGCAGG	GCCTacgatT
101	TCCGCGCCGA	TAAagcggcc	gGTGgctTTT	tcgGCataca	ggcgcaTatg
151	gCCTTTGTTT	ACCAgcatca	cgcggctgcg	accttgaTTT	TTGAACGATA
201	CTTCGCCgaT	GACAAATTCG	TCGGCTTGGT	ATTGCGCGGC	AACCTGCGCG
251	TATTTCAAAC	CGACAAAGCC	GATTTCGgga	ctggtaaACA	CCACGCCAAT
301	GGTgctgcgg	cGCAAACCGC	TGCCGATAtt	cgGtagcgg	ccccgcgta
351	ttgcccgcga	atcttacctt	ggtcggcggc	ttcatGCAGC	AGGGGCagtt
401	ggttggacgc	gtcgcccgcga	ataAAGATAT	GCGGAATgct	ggtCTGCATg
451	gtCAGCGGAT	CGGCAACGGG	tacgcgcgcg	gcgtctttgT	CGATATTGAT
501	GTTTTCCAAA	CCGATATtgT	CAACGTTTCGG	ACGGCgACCT	ACGGCTGCCA

```

551 ACATATATTC GGCAACAAAT ACGCCTTTTT CGCCATCCTG CTCCCAATGG
601 ACTtctACAT TGCCGTCTGC GTCGAGTTTG ACCTCGGTTT TAGCATCCAG
651 ATGCAGTTTC AATtctTCTC CGAACACGGC TTTCGCCTCG TCTGAAACAA
701 CGGGGTCGGA AATGCCGCCG ATGATTCCGC CCAAACCGAA AATTTCAACT
751 TTCACACCCA AACGGTGCAA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 14; ORF 004.ng>:

**g004 . pep**

```

1 MVERHIQHLR NGHLLHMRPC QQVSQMFGR AYDFRADKAA GGFFGIQAHM
51 AFVYQHAAAA TLIFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAN
101 GAAQAATAADI RVAAPRYCPA ILPWSAASCS RGSWLDASPA IKICGMLVCM
151 VSGSATGTPR ASLSILMFSK PILSTFGRRP TAANIYSATN TPFSPSCSQW
201 TSTLPSASSL TSVLASRCSF NSSPNTAFAS SETTGSEMP MIPPKPKIST
251 FPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 15>:

**m004 . seq**

```

1 ATGGTAGAAC GGCATATCCA GCATTGCGG AACGGTCATC TTCATTGAT
51 GTGCCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCAGG GCCTACGATT
101 TCCGCGCCGA TAAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTC ACCAGCATCA CGCGGCTGCG GCCTTGGTTT TTGAACGATA
201 CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTCAAGC CGACAAAGCC GATTGCGGA CTGGTAAACA CCACGCCGAT
301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGGCAATC TTGCCTTGGT CGGCAGCTTC ATGCAGCAGA GGCAGTTGGT
401 TGGACGCATC GCCTGCGATG AAGATATGCG GAATACTGGT CTGCATGGTC
451 AGCGGGTCGG CAACAGGTAC GCCGCGCGCA TCTTTTTTCA TATTGATATT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCATC GAGTTTGACC TCGGTTTTAG CATCCAGATG
651 CAGTTTCAAT TCTTCGCCGA ACACGGCGTT CGCCTCGTCT GAAACGACGG
701 GGTGCGAAAT GCCGCCGATG ATTCGCCCA AACCGAAAAT TTCAACTTTC
751 ACGCCCCAAC GGTGCAATGC CTGA

```

This corresponds to the amino acid sequence <SEQ ID 16; ORF 004>:

**m004 . pep**

```

1 MVERHIQHLR NGHLLHMCPS QQVRQMFGR AYDFRADKAA GGFFGIQAHM
51 AFVHQHAAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAD
101 GAAPQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPA KICGILVCMV
151 SGSATGTTPRA SFSILIFSKP ILSTFGRRP AASIYSATNT PFSPPSCSQWT
201 STLPSASSLT SVLASRCSFN SSPNTAFASS ETTGSEMP MIPPKPKISTF
251 TPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 17>:

**a004 . seq**

```

1 ATGGTAGAAC GGCATATCCA GCATTGCGG AACGGTCATC TTCATTGAT
51 GTGCCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCCCG ACCTACGATT
101 TCTGCGCCGA TGAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTT ACCAGCATCA CGCGGCTGCG GCCTTGGTTT TTGAACGATA
201 CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTCAAAC CGACAAAGCC GATTGCGGA CTGGTGAACA CTACGCCGAT
301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGGCAATC TTGCCTTGGT CGGCGGCTTC ATGCAGCAGG GGCAGTTGGT
401 TGGACGCGTC GCCCGCAATA AAGATATGCG GAATACTGGT CTGCATAGTC
451 AGCGGATCGG CAACGGGTAC GCCGCGCGCA TCTTTTTTCA TATTGATGTT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCTACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCGTC GAGTTTGACC TCGGTTTTAG CATCCAAATG
651 CAGTTTCAAT TCTTCACCGA ACACGGCTTT CGCCTCGTCT GAAACGACGG
701 GGTGCGAAAT GCCGCCGATG ATGCCACCA AACCGAAAAT TTCAACTTTC
751 ACGCCCCAAC GGTGCAATGC CTGA

```

This corresponds to the amino acid sequence <SEQ ID 18; ORF 004.a>:

a004.pep  
 1 MVERHIQHLR NGHLLHLMCPSS QQVRQMFGGGR TYDFCADEAA GGFFGIQAHM  
 51 AFVYQHHAHA ALVFERFYFAD DKFVGLVLRG NLRVFQTDKA DLRTGEHYAD  
 101 GAAQAQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAI KICGILVCIV  
 151 SGSATGTPRA SFSILMFSKP ILSTFGRRPT AASIYSATNT PFSPSCSQWT  
 201 STLPSASSLA SVLASKCSFN SSPNTAFASS ETTGSEMPPM MPPKPKISTF  
 251 TPKRCNA\*

m004/a004 94.9% identity over a 257 aa overlap

m004.pep	10	20	30	40	50	60
	MVERHIQHLRNGHLHLMCPSSQQVRQMFGGGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA					
a004	MVERHIQHLRNGHLHLMCPSSQQVRQMFGGRTYDFCADEAAAGGFFGIQAHMAFVYQHHAHA					
	10	20	30	40	50	60
m004.pep	70	80	90	100	110	120
	ALVFERFYFADDKFVGLVLRGNLRFQTDKADLRTGKHHADGAAPQTAADIRVAAALSPAI					
a004	ALVFERFYFADDKFVGLVLRGNLRFQTDKADLRTGEHYADGAAQAQTAADIRVAAALSPAI					
	70	80	90	100	110	120
m004.pep	130	140	150	160	170	180
	LPWSAASCSRGSWLDASPAIKICGILVCMVSGSATGTPRASFSILIFSKPILSTFGRRPT					
a004	LPWSAASCSRGSWLDASPAIKICGILVCIVSGSATGTPRASFSILMFSKPILSTFGRRPT					
	130	140	150	160	170	180
m004.pep	190	200	210	220	230	240
	AASIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMPPM					
a004	AASIYSATNTPFSPSCSQWTSTLPSASSLASVLASKCSFNSSPNTAFASSETTGSEMPPM					
	190	200	210	220	230	240
m004.pep	250					
	IPPKPKISTFTPKRCNAX					
a004	MPPKPKISTFTPKRCNAX					
	250					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 004 shows 93.4% identity over a 258 aa overlap with a predicted ORF (ORF 004.ng) from *N. gonorrhoeae*:

m004/g004

m004.pep	10	20	30	40	50	60
	MVERHIQHLRNGHLHLMCPSSQQVRQMFGGGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA					
g004	MVERHIQHLRNGHLHLMRQCQVVSQMFGGGRAYDFRADKAAGGFFGIQAHMAFVYQHHAHA					
	10	20	30	40	50	60
m004.pep	70	80	90	100	110	119
	ALVFERFYFADDKFVGLVLRGNLRFQTDKADLRTGKHHADGAAPQTAADIRVAAA-LSPA					
g004	TLIFERYFADDKFVGLVLRGNLRFQTDKADLRTGKHHANGAAQAQTAADIRVAAPRYCPA					
	70	80	90	100	110	120
m004.pep	120	130	140	150	160	179
	ILPWSAASCSRGSWLDASPAIKICGILVCMVSGSATGTPRASFSILIFSKPILSTFGRRP					



```

g004      |||||:||||:|||||:||||:|||||
          ILPWSAASCSRGSWLDASPAIKICGMLVCMVSGSATGTPRASLSILMFSPILSTFGRRP
          130      140      150      160      170      180

m004.pep  180      190      200      210      220      230      239
          TAASIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMP
          |||||:|||||:|||||:|||||:|||||:|||||:|||||
g004      TAANIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMP
          190      200      210      220      230      240

m004.pep  240      250
          MIPPKPKISTFTPKRCNAX
          |||||
g004      MIPPKPKISTFTPKRCNA
          250

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 19>:

```

g005.seq
1  ATGGGGATGG ACAATATTGA TATGTTTCATG CCTGAACAAG AGGAAATCCA
51 ATCAATGTGG AAAGAAATTT TACTGAATTA CGGTATTTTC CTGCTCGAAC
101 TGCTTACCGT GTTCGGCGCA ATTGCGCTGA TTGTGTTGGC TATCGTACAG
151 AGTAAGAAAC AGTCGGAAG CCGCAGTGTC GTACTGACAG ATTTTTCGGA
201 AAATTATAAA AAACAGCGGC AATCGTTTGA AACATTCTTT TTAAGCGAGG
251 AAGAGACAAA ACATCAGGAA AAAAAAGAAA AGAAAAAGGA AAAGGCGGAA
301 GCCAAAGCAG AGAAAAAGCG TTTGAAGGAG GGCGGGGAGA AATCTGCCGA
351 AACGCAAAA TCCCGCCTTT TTGTGTTGGA TTTTGACGGC GATTTGTATG
401 CACACGCCGT AGAATCCTTG CGTCATGAGA TTACGGCGGT GCTTTTGATT
451 GCCAAGCCTG AAGATGAGGT TCTGCTCAGA TTGGAAAGTC CGGGCGGCGT
501 GGTTACGGT TACGGTTTGG CGGCTTCGCA GCTTAGGCGT TTGCGCGAAC
551 GCAATATTCC GCTGAccgtc gccgTCGATA AGGTTCGCGC AAGCGgccgc
601 tatatgatgg cgtgtgtgGC GGATAAAATT GTTCCCGTc cgtttgcggt
651 catcggttcg gtgggtgtgg tgGcggaagt gcCGAATATC CAccgCctGT
701 TGAAAAACA TGATATTGAT GTGGATGTGA TGACGGCGGG CGAATTTAAG
751 CGCACGGTTA CTTTTATGGG TGAAAATACG GAAAAGGGCA AACAGAAATT
801 CCGGCAGGAA CTGGAGGAAA CGCATCAGTT GTTCAAGCAG TTTGTCAGTG
851 AAAACCGCCC CGGGTTGGAT ATTGAAAAA TAGCGACGGG CGAGCATTGG
901 TTCGGCCGGC AGGCGTTGGC GTTGAAGTTG ATTGACGAGA TTTCGACCAG
951 TGATGATTTG TTGTTGAAAG CGTTTGAAAA CAAACAGGtt aTCGAAGTGA
1001 AATATCAGGA GAAGCGAAGC CTGATCCAGC GCATTGGTTT GCAGGCGGAA
1051 GCTTCCGTTG AAAAGTTGTT TGCCAAACTT GTCAACCGGC GAGCGGATGT
1101 GATGTAG

```

This corresponds to the amino acid sequence <SEQ ID 20; ORF 005.ng>:

```

g005.pep
1  MGMDNIDMF PEQEEIQSMW KEILLNYGIF LLELLTVFGA IALIVLAIVQ
51 SKKQESGVS VLTDFSENYK KQRQSFETFF LSEEETKHQE KKEKKKEKAE
101 AKAEEKRLKE GGEKSAETQK SRLFVLDFDG DLYAHAVESL RHEITAVLLI
151 AKPEDEVLLR LESPGGVVHG YGLAASQLRR LRERNIPLTV AVDKVAASGG
201 YMMACVADKI VSAPFAVIGS VGVVAEVPNI HRLKKHDID VDVMTAGEFK
251 RTVTFMGENT EKGKQKFRQE LEETHQLFKQ FVSENRPGLD IEKIATGEHW
301 FGRQALALNL IDEISTSDDL LLKAFENKQV IEVKYQEKRS LIQIRIGLQAE
351 ASVEKLFACL VNRRADV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 21>:

```

m005.seq
1  ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51 GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
151 AACAGTCGG AWAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
201 TAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
251 CACAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGGC GGAAGCCAAA

```

```

301 GCAGAGAAAA A.CGTTTGAA GGAGGGTGGG GAGAAATCTG CCGAAACGCA
351 nAAATCACGC CTTTTGTGT TGGANNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGCGAGCGG CGGTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGCTTCC GCTCCGTTTG CGATTGTCCG
651 TTCGGTGGGT GTGGTGGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAGAG GGCAACAGA AATCCGACA
801 GGAAGTGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTATTCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTGCAGGC GGAAGCTTCT
1051 GTTGAAAAGT TGTTTGCCAA ACTTGTCAAC CGGCGGGCGG ATGTGATGT A
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 22; ORF 005>:

## m005.pep

```

1 MDNIDMFMP EEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQXSGSVVL TDFSENYKKQ RQSFEAFFLS GEEAQHQEKE EKKKEKAEAK
101 AEKXRLKEGG EKSAETXKSR LFVLXXXXXX XXXXXXXXXXX XXXXXXXXXXX
151 XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXASGGYM
201 MACVADKIAS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG
301 RQALALNLID EISTSDDLLL KAFENKQVIE VKYQEKQSLI QRIQLQAEAS
351 VEKLEAKLVN RRADV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 23>:

## a005.seq

```

1 ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51 GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
151 AAACAGTCGG AAAGCGGCAG TGTCGTACTG ACGGATTTT CCGAAAATTA
201 TAAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
251 CAAAACATCA GGAAAAAGAG GAAAGAAAA AGGAAAAGGC GGAAGCCAAA
301 GCAGAGAAAA AGCGTTTGAA GGAGGGTGGG GAGAAATCTT CCGAAACGCA
351 AAAATCCCGC CTTTTGTGT TGGATTTTGA CGGCGATTG TATGCACACG
401 CCGTAGAATC CTGCGTCAT GAGATTACGG CGGTGCTTTT GATTGCCAAG
451 CCTGAAGATG AGGTTCTGCT TAGATTGGAA AGTCCGGGCG GCGTGGTTCA
501 CGGTTACGGT TTGGCGGCTT CGCAGCTTAG GCGTTTGCGC GAACGCAATA
551 TTCCGCTGAC CGTCGCGTC GATAAGGTGG CGGCGAGCGG TGTTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGTTTCC GCTCCGTTTG CGATTGTCCG
651 TTCGGTGGGT GTTGTAGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAGAG GGCAACAGA AATCCGACA
801 GGAAGTGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTATTCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTGCAGGC GGAAGCTTCT
1051 GTTGAAAAGT TGTTTGCCAA ACTTGTCAAC CGGCGGGCGG ATGTGATGTA
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 24; ORF 005.a>:

## a005.pep

```

1 MDNIDMFMP EEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQSESGSVVL TDFSENYKKQ RQSFEAFFLS GEEAKHQEKE EKKKEKAEAK
101 AEKXRLKEGG EKSETQKSR LFVLDLFDGDL YAHAVESLRH EITAVLLIAK
151 PEDEVLLRLE SPGGVVHGYG LAASQLRRLR ERNIPLTVAV DKVAASGGYM
201 MACVADKIVS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG

```

301 RQALALNLID EISTSDDLLL KAFENKQVIE VKYQEKQSLI QRIQLQAEAS  
 351 VEKLFKLVN RRADVM\*

m005/a005 79.2% identity over a 366 aa overlap

m005.pep	MDNIDMFMPQEIEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQXSXSGSVVL
a005	MDNIDMFMPQEIEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSESGSVVL
m005.pep	TDFSENYKKQSQSFEEAFLSGEEAQHQEKEKKKEKAEAKAEKKRLKEGGEKSAETXKSR
a005	TDFSENYKKQSQSFEEAFLSGEEAKHQEKEKKKEKAEAKAEKKRLKEGGEKSSETQKSR
m005.pep	LFVLXX
a005	LFVLDFDGLYAHAVESLRHEITAVLLIAKPEDEVLLRLESPPGVVHGYGLAASQLRRLR
m005.pep	XXXXXXXXXXXXXXXXASGGYMMACVADKIASAPFAIVGSGVVAEVPNIHRLKKHDIDVD
a005	ERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAIVGSGVVAEVPNIHRLKKHDIDVD
m005.pep	VMTAGEFKRTVTFMGENTKQKQFRQELEETHQLFKQFVSENRPQLDIEEVATGEHWFG
a005	VMTAGEFKRTVTFMGENTKQKQFRQELEETHQLFKQFVSENRPQLDIEEVATGEHWFG
m005.pep	RQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFKLVN
a005	RQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFKLVN
m005.pep	RRADVMX
a005	RRADVMX

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 005 shows 77.0% identity over a 366 aa overlap with a predicted ORF (ORF 005.ng) from *N. gonorrhoeae*:

m005/g005

m005.pep	MDNIDMFMPQEIEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQXSXSGSV
g005	MGMDNIDMFMPQEIEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSESGSV
m005.pep	VLTDFSENYKKQSQSFEEAFLSGEEAQHQEKEKKKEKAEAKAEKKRLKEGGEKSAETXK
g005	VLTDFSENYKKQSQSFETFFLSEETKHQEKKEKKKEKAEAKAEKKRLKEGGEKSAETQK

	120	130	140	150	160	170
m005.pep	SRLFVLXX					
			:			
g005	SRLFVLDFDGLYAHAVESLRHEITAVLLIAKPEDEVLLRLESPPGGVVHGYGLAASQLRR					
	130	140	150	160	170	180
	180	190	200	210	220	230
m005.pep	XXXXXXXXXXXXXXXXASGGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRLKKHDID					
	:					
g005	LRERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAVIGSVGVVAEVPNIHRLKKHDID					
	190	200	210	220	230	240
	240	250	260	270	280	290
m005.pep	VDVMTAGEFKRTVTFMGENTEKQKQFRQLEETHQLFKQFVSENRPQLDIEEVATGEHW					
g005	VDVMTAGEFKRTVTFMGENTEKQKQFRQLEETHQLFKQFVSENRPGLDIEKIATGEHW					
	250	260	270	280	290	300
	300	310	320	330	340	350
m005.pep	FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFACL					
g005	FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKRSLIQIRIGLQAEASVEKLFACL					
	310	320	330	340	350	360
	360					
m005.pep	VNRRADVMX					
g005	VNRRADVMX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 25>:

## g006.seq

```

1   ATGCTGCTGG TGCTggaatt ttggttCGGc gtGtCGGCGG TGGGCatact
51  tgCGTTGTTT TTATGGCttt TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCCG CCTGAACAAC AGCTTGGAAC gcgACAACCA CTTTATCCGA
151 AAAGGCGACG AGCGGCAGCT GTACCGCCAT TACGGACTGG TTTCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCG
251 CGGCGATGGG TATTTTGTTT GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACGGCA GCGCGGGGCA TATTTATTCG GTCGGCACTT ATCTGTGGAT
351 GTTTGCCATG AGTTTGGACG ATGTGCCGCG ATTGGTCGAA CAATATTCCA
401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCCGAACG GAACATCAAA
451 GCCGGAAGT GA

```

This corresponds to the amino acid sequence <SEQ ID 26; ORF 006.ng>:

## g006.pep

```

1   MLLVLEFWFG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
51  KGDERQLYRH YGLVSRLRVL ISNREAFGYL CVGAAMGILF GFAFVMMTLK
101 GYGSAGHIYS VGTYLWMFAM SLDDVPRLVE QYSLNKDIGQ RIEWSERNIK
151 AGT*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 27>:

## m006.seq

```

1   ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT
51  TGCGTTGTTT TTATGGCttt TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCCG CCTGAACAAC AGCTTGGAAC GCGACAACCA CTTTATCCGA
151 AAAGGCGACC GCGGCAGCT GTACCGCCAT TACGGACTGC TTTCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCA
251 CGGCGATGGG TATTTTGTTT GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACAGCA GCGCGGGGCA TGTCTATTCG GTCGGCACTT ATCTGTGGAT

```

351 GTTTGCCATG AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA  
 401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTGCGAACG GAACATCAAA  
 451 GCCGGAACCTGA

This corresponds to the amino acid sequence <SEQ ID 28; ORF 006>:

**m006.pep**

1 MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR  
 51 KGDRRQLYRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK  
 101 GYSSAGHVYS VGTYLWMFAM SLDDVPRLVE QYSNLKDIGQ RIEWSERNIK  
 151 AGT\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 29>:

**a006.seq**

1 ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT  
 51 TCGGTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC  
 101 TGTATTTCCT CCTGAAGAAC AGCTTGAAC GCGACAACCA CTTTATCCGA  
 151 AAAGGCGACG AGCGGCAGCT GGACCGCCAT TACGGACTGC TTGCGCGCCT  
 201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TCGCTCGGCA  
 251 CGCGCATGGG TATTTTGTTT GGCTTTGCTT TTGTGATGAT GACGCTCAAA  
 301 GGCTACAGCA GCGCGGGGCA TGTCTATTCT GTCGGCACTT ATCTGTGGAT  
 351 GTTTGCCATA AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA  
 401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTGCGAACG GAACATCAAA  
 451 GCCGGAACCT GA

This corresponds to the amino acid sequence <SEQ ID 30; ORF 006.a>:

**a006.pep**

1 MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLKN SLERDNHFIR  
 51 KGDERQLDRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK  
 101 GYSSAGHVYS VGTYLWMFAI SLDDVPRLVE QYSNLKDIGQ RIEWSKRNIK  
 151 AGT\*

**m006/a006** 96.7% identity over a 153 aa overlap

	10	20	30	40	50	60
m006.pep	MLLVLEFWVGVS AVGILALFLWLLPRFAAISENLYFRLNNSLERDNHFIRKGDRRQLYRH					
	:     :     :					
a006	MLLVLEFWVGVS AVGILALFLWLLPRFAAISENLYFRLKNSLERDNHFIRKGDERQLDRH					
	10	20	30	40	50	60
	70	80	90	100	110	120
m006.pep	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAM					
	:     :     :     :     :					
a006	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAI					
	70	80	90	100	110	120
	130	140	150			
m006.pep	SLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX					
	:     :					
a006	SLDDVPRLVEQYSNLKDIGQRIEWSKRNIKAGTX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 006 shows 95.4% identity over a 153 aa overlap with a predicted ORF (ORF 006.ng) from *N. gonorrhoeae*:

**m006/g006**

	10	20	30	40	50	60
m006.pep	MLLVLEFWVGVS AVGILALFLWLLPRFAAISENLYFRLNNSLERDNHFIRKGDRRQLYRH					

[illegible]

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 31>:

```
g006-1.seq
1  ATGTGGA AAA TGTTGAAACA CATAGCCAAA ACCCACC GCA AGCGATTGAT
51  TGGCACATTT TCCCCGGTCG GACTGGAAAA CCTTTTGATG CTGGGGTATC
101 CGGTGTTTGG CGGCTGGGCG ATTAATGCCG TGATTGCGGG GAGGGTGTGG
151 CAGGCGTTGC TGTACGCTTT GGTGTATTTT TTGATGTGGC TGGTCCGGTG
201 GGCACGCGCG ATTGCCGATA CGCGCACGTT TACGCGGATT TATACCGAAA
251 TCGCCGTGCC GGTTGTGTTG GAACAACGGC AGCGGCAAGT CCCGCATTCA
301 GCGGTAAC TG CACGGGTTGC CCTGTCGCGT GAATTTGTCA GCTTTTTTGA
351 AGAACACCTG CCGATTGCCG CGACATCCGT CGTATCCATA TTCGCGCGGT
401 GCATCATGCT GCTGGTGCTG GAATTTTGGG TCGCGCGTGC GCGGGTGGCG
451 ATACTGCGT TGTTTTTATG GCTTTTGCCA CGTTTGTCCG CCATCAGCGA
501 AAACCTGTAT TTCCGCGCTA ACAACAGCTT GGAACGCGAC AACCACTTTA
551 TCCGAAAAAG CGACGAGCGG CAGCTGTACC GCCATTACGG ACTGGTTTCG
601 CGCTGCGTG TGCTGATTTT CAAACCGGAA GCCTTCGGCT ATCTCTGCGT
651 CGGCGCGGCG ATGGGTATTT TGTTCCGCTT TGCTTTTGTG ATGATCAGCG
701 TCAAAGGCTA CGGCAGCGCG GGGCATATTT ATTCGGTCGG CACTTATCTG
751 TGGATGTTTG CCATGAGTTT GGACGATGTG CCGCGATTGG TCGAACATA
801 TTCCAATTTG AAAGCATCG GACAACGGAT AGAGTGGTCG GAACGGAA
851 TCAAAGCCGG AACTGA
```

This corresponds to the amino acid sequence <SEQ ID 32; ORF 006-1.ng>:

g006-1.pep

1	MWKMLKHIAK	THRKRLIGTF	SPVGLENLLM	LGYPVFGGWA	INAVIAGRVM
51	<u>QALLYALVVF</u>	<u>LMWLVGARR</u>	IADTRTFTRI	YTEIAVPVVL	EQRQRQVPHS
101	AVTARVALSR	EFVSFFEEHL	PIAATSVVSI	FGACIMLLVL	<u>EFWVGVSavg</u>
151	<u>ILALFWLLP</u>	RFAAISENLY	FRLLNSLERD	NHFIRKGDGR	QLYRHYGLVS
201	RLRVLISNRE	AFGYLCVGAA	MGLNFGFAFV	MMTLKGYGSA	GHIYSVGTYL
251	WMFAMSLDDV	PRLVEQYSNL	KDIGQRIEWS	ERNIKAGT*	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 33>:

```
m006-1.seq
1  ATGTGGA AAA TGT TGA AACA CATAGCCCAA ACCCACC GCA AGCGATTGAT
51  TGGCACATTT TCCCTGGT CG GACTGGA AAA CCTTTTGATG CTGGTGTATC
101 CGGTGTTTGG CGGCCGG CG ATCAATGCCG TGATTGCGGG GGAGGTGTGG
151 CAGGCGTTGC TGTACGCTTT GGTGTGCTTT TTGATGTGGC TGCTCGGTGC
201 GGTGCGGCGG ATTGCCGATA CGCGCACGTT TACGCGGATT TATACCGAAA
251 TCGCCGTGCC GGTGCTGTTG GAACAGCGCG AGCGACAAGT CCCGCATT CG
301 GCGGTA ACTG CGCGGGTTGC CCTGTCGCGT GATTTGTCA GCTTTTTTGA
351 AGAACACTG CCGATTGCCG CCGATTCCGT CAGATCCATA TTCGGCGCGT
401 GCATCATGCT GCTGGTGCTG GAATTTTGGG TCGGCGGTGC GCGCGTGGCG
451 ATACTTGC GTTTTTTATG GCTTTTGCCA CGTTTGTCCG CCATCAGCGA
501 AAACCTGTAT TTCCGCCTGA ACAACAGCTT GGAACGCGCA AACCCTTTA
551 TCCGAAAAGG CGACCGGCGG CAGCTGTACC GCCATTACGG ACTGCTTGCG
601 CGCCTGCGTG TGCTGATTTT CAACCGCGAA GCCTTCGGCT ATCTCTGCGT
```

m006-1.ppt

1	MWKMLKHIAQ	THRKRLIGTF	SLVGLENLLM	LVYPVFGGGR	INAVIAGEVW
51	QALLYALVVL	LMWLVGAVRR	IADTRTFTRI	YTEIAVPVVL	EQRQRQVPHS
101	AVTARVALSR	EFVSFFEEHL	PIATSVVSI	FGACIMLLVL	EFWVGVSavg
151	ILALFLWLLP	RFAAISENLY	FRLNNSLDR	NHFIRKGGRR	QLYRHYGLLA
201	RLRVLIISRE	AFGYLCVGTA	MGILEGFADFV	MMTLKGYSR	GHVYSVGTYL
251	WMFAMSLDDV	PRLVEQYSNL	KDIGORIEWS	ERNIKAGT*	

[illegible]

a006-1.seq (partial)

1	..AGCCAAAACC	ACCGCAAGCG	ATTGATTGGC	ACATTTTTTC	TGGTCGGA
51	GGAAAACCTT	TTGATGCTGG	TGTATCCGGT	GTTTGGCGGC	TGGGCGATTA
101	ATGCCGTGAT	TGCGGGGCGG	GCGTGGCAGG	CGTTGCTGTA	CGCTTTGGTT
151	TGTGCTTTGA	TGTGGCTGGT	CGGTGCGGCG	CGGCGGATTG	CCGATACCGG
201	CACGTTTACG	CGGATTTATA	CCGAAATCGC	CGTGCCGGTT	GTGTTGGAAC
251	AGCGGCAGCG	GCAAGTCCCG	CATTGCGCGG	TAAC TGCGCG	GGTTGCCCTG
301	TCCGCTGAGT	TTGTCACTTC	TTT TGAAGAA	CACCTGCCGA	TTGCCGCGAC
351	ATCCGTCGTA	TCCATATTTC	GCGCGTGCAT	CATGCTGCTG	GTGCTGGAAT
401	TTTGGGTCGG	CGTGTGCGCG	TGGGCGATAC	TTGCGTTGTT	TTTATGGCTT
451	TTGCCACGTT	TTGCCGCCAT	CAGCGAAAC	CTGTATTTCC	GCGCTGAAGAA
501	CAGCTTGGAA	CGCGACAACC	ACTTTATCCG	AAAAGGCGAC	GAGCGGCAGC
551	TGGACCGCCA	TTACGGAATG	CTTGCGCGCC	TGCGCTGTGCT	GATTTCCAAC
601	CGCGAAGCCT	TGCGCTACTC	TGCGCTGCGC	ACGGCGATGG	GTATTTTGTT
651	CGGCTTTGCT	TTTGTGATGA	TGACGCTCAA	AGGCTACAGC	AGCGCGGGGG

701 ATGTCTATTC GGTCGGCACT TATCTGTGGA TGTTTGCCAT AAGTTTGGAC  
 751 GACGTGCCGC GATTGGTCGA ACAATATTCC AATTTGAAAG ACATCGGACA  
 801 ACGGATAGAG TGGTCGAAAC GGAACATCAA AGCCGGAAC TGA

This corresponds to the amino acid sequence <SEQ ID 36; ORF 006-1.a>:

a006-1.pep (partial)

1 ..SQNHRKRLIG TFFLVGLENL LMLVYPVFGG WAINAVIAGO AWOALLYALV  
 51 VLLMWLVGAA RRIADTRTFT RIYTEIAVPV VLEQRQRQVP HSAVTARVAL  
 101 SREFVSFFEE HLPAAATSVV SIFGACIMLL VLEFWVGUSA VGILALFLWL  
 151 LPRFAAISEN LYFRLKNSLE RDNHFIRKGD ERQLDRHYGL LARLRVLISN  
 201 REAFGYLCVG TAMGILFGFA FVMMTLKGYS SAGHVYSVGT YLWMFAISLD  
 251 DVPRLVEQYS NLKDIGQRIE WSKRNIKAGT \*

a006-1/m006-1 95.7% identity in 280 aa overlap

	10	20	30	40	50
a006-1.pep	SQNHRKRLIGTFFLVGLENLLMLVYPVFGGWAINAVIAGOAWOALLYALVVL				
m006-1	MWKMLKHIAQTHRKRLLIGTFFSLVGLENLLMLVYPVFGGRAINAVIAGEVWQALLYALVVL				
	10	20	30	40	50
	60	70	80	90	100
a006-1.pep	LMWLVGAAARRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL				
m006-1	LMWLVGAVRRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL				
	70	80	90	100	110
	120	130	140	150	160
a006-1.pep	PIAATSVVSIFGACIMLLVLEFWVGVSavgilALFLWLLPRFAAISENLYFRLKNSLERD				
m006-1	PIAATSVVSIFGACIMLLVLEFWVGVSavgilALFLWLLPRFAAISENLYFRLNNSLERD				
	130	140	150	160	170
	180	190	200	210	220
a006-1.pep	NHFIRKGDERRQLDRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYS				
m006-1	NHFIRKGDERRQLDRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYS				
	190	200	210	220	230
	240	250	260	270	280
a006-1.pep	GHVYSVGTYLWMFAISLDDVPRLVEQYSNLKDIGQRIEWSKRNIKAGTX				
m006-1	GHVYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX				
	250	260	270	280	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 37>:

g007.seq

1 atgaACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGcgC  
 51 CGCcGCTTCT GCCGccgaca acAGCatcat gaCaAAAGGG CAAAAAGTGT  
 101 ACGAATCcAa ctGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC  
 151 ACTGCGtTTC CTccgctTTT CCggtcgGac tgtattatga acaAACCGCa  
 201 cgTCctgctg cacagcatgg tcaaaggcAt cgacgggaca ttcaaagtgg  
 251 agcggcaaaa cctacgacgg atttatgCcc gcaaccgcca tcagcgATGC  
 301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

This corresponds to the amino acid sequence <SEQ ID 38; ORF 007.ng>:

g007.pep

1 MNTRLPTAF ILCCLCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG  
 51 TAFPPLFRSD CIMNKPHVLL HSMVKGIDGT FKVERQNLRR IYARNRHQRC  
 101 GHCRRRHLYH ERL\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 39>:



## m007.seq

```

1   ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC
51  CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251 AACGGCAAAA CCTACAACGG ATTCATGCCC GCAACGCCA TCAGCGATGC
301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

```

This corresponds to the amino acid sequence <SEQ ID 40; ORF 007>:

## m007.pep

```

1   MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51  TMFPPLYRSD FIMKKPQVLL HSMVKINGT IKVXRQNLQR IHARNRHQRC
101 GHCRRRHLYH ERL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 41>:

## a007.seq

```

1   ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC
51  CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251 AACGGCAAAA CCTACAACGG ATTCATGCCC GCCACTGCCA TCAGCGATGC
301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

```

This corresponds to the amino acid sequence <SEQ ID 42; ORF 007.a>:

## a007.pep

```

1   MNTTRLPTAL VLGCCLAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51  TMFPPLYRSD FIMKKPQVLL HSMVKINGT IKVXRQNLQR IHARHCHQRC
101 GHCRRRHLYH ERL*

```

m007/a007 97.3% identity over a 113 aa overlap

	10	20	30	40	50	60
m007.pep	MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
	:					
a007	MNTTRLPTALVLGCCLAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
	10	20	30	40	50	60
	70	80	90	100	110	
m007.pep	FIMKKPQVLLHSMVKINGTIKVXRQNLQRIHARNRHQRCGHCRRRHLYHERLX					
	:					
a007	FIMKKPQVLLHSMVKINGTIKVXRQNLQRIHARHCHQRCGHCRRRHLYHERLX					
	70	80	90	100	110	

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 007 shows 86.7% identity over a 113 aa overlap with a predicted ORF (ORF 007.ng) from *N. gonorrhoeae*:

## m007/g007

	10	20	30	40	50	60
m007.pep	MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
	: :					
g007	MNTTRLPTAFILCCLCAAASAADNSIMTKGQKVYESNCIACHGKKGEGRGTAFFPLFRSD					
	10	20	30	40	50	60
	70	80	90	100	110	
m007.pep	FIMKKPQVLLHSMVKINGTIKVXRQNLQRIHARNRHQRCGHCRRRHLYHERLX					
	:     :					

CIMNKPHVLLHSMVKGIDGTFKVERQNLRRRIYARNRHQRCGHCRRRHLYHERL  
70 80 90 100 110

g007-1.seq (partial)

1	ATGAACACAA	CCCGACTGCC	GACCGCCTTC	ATCTTGTGCT	GCCTCTGCGC
51	CGCCGCTTCT	GCCGCCGACA	ACAGCATCAT	GACAAAAGGG	CAAAAAGTGT
101	ACGAATCCAA	CTGCATCGCC	TGCCACGGCA	AGAAAGGGGA	AGGGCGCGGC
151	ACTCGCTTCT	CTCCGCTTTT	CCGTCGCGAC	TATATTATGA	ACAAACCGCA
201	CGTCGTCGTC	CACAGCATGG	TCAAAGGCAT	CAACGGTACA	ATCAAAGTCA
251	ACGGCAAAAC	CTACAAACGGA	TTATATGCCG	CAACCGCCAT	CAGCGATGCG
301	GACATTGCCG	CCGTCGCCAC	TTATATCATG	AACGCTTTTG	ACAACGGCGG
351	CGGAAGCGTT	ACCGAAAAAG	ACGTAAACA	GGCAAAGGC	AAAAAAAAC

g007-1.pep (partial)

1 MNTTRLPTAF ILCCCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG  
51 TAFPPLFRSD YIMNKPVL L HSMVKGINGT IKVNGKTYNG FMPATAISDA  
101 DIAAVATYIM NAFDNGGGSV TEKDVQKQAKG KKN...

m007-1.seq

1	ATGAACACAA	CCCGACTGCC	GACCGCCCTC	GTCTTGGGCT	GCTTCTGCGC
51	CGCCGCTTCT	GCCGCCGACA	ACAGCATCAT	GACAAAAGGG	CAAAAAGTGT
101	ACGAATCCAA	CTGCGTCGCC	TGCCACGGCA	AAAAGGGCGA	AGGCCGCGGA
151	ACCATGTTTC	CGCCGCTCTA	CCGCTCCGAC	TTCTATCATG	AAAAACCGCA
201	GGTGCTGCTG	CACAGCATGG	TCAAAGGCAT	CAACGGTACA	ATCAAAGTCA
251	ACGGCAAAAC	CTACAACGGA	TTCATGCCCG	CAACGCCCAT	CCGCATGCGC
301	GACATTGCCG	CCGTCGCCAC	TTATATCATG	AACGCCTTTG	ACAACGGCGG
351	CGGAAGCGTT	ACCGAAAAAG	ACGTAAAAACA	GGCAAAAAGC	AAAAAAAAT
401	AA				

m007-1.pgp

1 MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG  
51 TMFPPLYRSD FIMKKPOVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA  
101 DIAAVATYIM NAFDNGGGSV TEKDVKOAKS KKN\*

m007-1 / g007-1 91.7% identity in 133 aa overlap

		10	20	30	40	50	60
m007-1.pep	MN	TTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKKGEGRGTMFPPLYRSD					
	:	:	:	:	:	:	:
g007-1	MN	TTRLPTAFILCCLCAAASAADNSIMTKGQKVYESNCIACHGKKKGEGRGTAFFPLFRSD					
	:	:	:	:	:	:	:
		10	20	30	40	50	60
		70	80	90	100	110	120
m007-1.pep	FIMK	KPQVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV					
	:	:	:	:	:	:	:
g007-1	YIMN	KPHVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV					
	:	:	:	:	:	:	:
		70	80	90	100	110	120
		130					
m007-1.pep	TEKD	VKQA	SKKNX				
	:	:	:	:	:	:	:
g007-1	TEKD	VKQA	GKKNN				
	:	:	:	:	:	:	:
		130					

a007-1.seq (partial)

1 ATGAAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC  
51 CGCCGCTTCT GCGCGCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT  
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGCGGA AGGCCGCGGA

```

151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACC GCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CCACTGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAAG ACGTAAACA GGCAAAAAAC AAAAAA..

```

This corresponds to the amino acid sequence <SEQ ID 48; ORF 007-1.a>:

**a007-1.pep (partial)**

```

1 MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRSD FIMKKPVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKN KK..

```

**m007-1/a007-1 98.5% identity in 132 aa overlap**

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALVLGCLCAAAS	AADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD			
a007-1	MNTTRLPTALVLGCLCAAAS	AADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD			
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPVLLHSMVKGINGT	IKVNGKTYNGFMPATAISDADIAAVATYIM	NAFDNGGGSV			
a007-1	FIMKKPVLLHSMVKGINGT	IKVNGKTYNGFMPATAISDADIAAVATYIM	NAFDNGGGSV			
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVQAKSKKNX					
a007-1	TEKDVQAKNKK					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 49>:

**g008.seq**

```

1 ATGAACAACA GACATTTTGC CGTCatcgCC TTGGGCAGCA ACCTTGACAA
51 CCCCGCACAA CAAATacgcg gcgcattaga cgcgctctcg tcccatcctg
101 acatccggct tgaaCagggt tcctcactgt aTatgaccgc acctgtcgggt
151 tacgAcaaTC agcccgATTT CATCaatgcc gTCTgcaccg TTTCCACCAC
201 CtTGGACGGC ATTGcccTGC TTGCCgaACT CAAccgTATC GAAGCCGATT
251 TCGGACGCGA aCGCAGTTTC CGCAATGCAC CGCGCACATT GGATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGGCC TTACCCTGCC
351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATAcGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAATACG GAAAGGTTGT CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGACA GGTA

```

This corresponds to the amino acid sequence <SEQ ID 50; ORF 008.ng>:

**g008.pep**

```

1 MNNRHFAVIA LGSNLDNPAQ QIRGALDALS SHPDIRLEQV SSLYMTAPVG
51 YDNQPDFINA VCTVSTTLDG IALLAELNRI EADFGRRERF RNAPRTLDDLD
101 IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKYGKVVELS
151 KRLGNQGIRL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 51>:

**m008.seq**

```

1 ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
51 CCCTGCTCAA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTT TGTCaATGCC GTCTGCACCG TTTCCACCAC
201 TCTGGACGGC ATTGCCyTGC TTGCCgaACT CAACCGTATC GAGGCTGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GkATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACACsCGAC TcACcTGCC

```

```

351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATCCGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTGTTTTA GGAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGyTGG GCAATCAAGG TATCCGTCTT TTACCGGACA GGTAAAT

```

This corresponds to the amino acid sequence <SEQ ID 52; ORF 008>:

**m008.pep**

```

1 MNNRHFVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
51 YDNQPDFVNA VCTVSTTLTG IALLAELNRI EADFGRRERF RNAPRTLXLD
101 IIDFDGISSD DTRLTLPHPR AHERSFVIRP LAEILPDFVL GKHKQVAELS
151 KRLGNQGIRL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 53>:

**a008.seq**

```

1 ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
51 CCCTGCCCAA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTT CGTCAATGCC GTCTGCACCG TTTCCACCAC
201 CTTGGACGGC ATTGCCCTGC TTGCCGAAT CAACCGTATC GAAGCCGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GGATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGAC TCACCTGCC
351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CACACGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGATA AGTAA

```

This corresponds to the amino acid sequence <SEQ ID 54; ORF 008.a>:

**a008.pep**

```

1 MNNRHFVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
51 YDNQPDFVNA VCTVSTTLTG IALLAELNRI EADFGRRERF RNAPRTLXLD
101 IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKHKQVAELS
151 KRLGNQGIRL LPDK*

```

**m008/a008** 97.6% identity over a 164 aa overlap

m008.pep	10	20	30	40	50	60
	MNNRHFVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
a008	MNNRHFVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
	10	20	30	40	50	60
m008.pep	70	80	90	100	110	120
	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
a008	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDPRLTLPHPR					
	70	80	90	100	110	120
m008.pep	130	140	150	160		
	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX					
a008	AHERSFVIRPLAEILPDFILGKHGKVAELSKRLGNQGIRLLPDKX					
	130	140	150	160		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 008 shows 92.7% identity over a 164 aa overlap with a predicted ORF (ORF008.ng) from *N. gonorrhoeae*:

**m008/g008**

m008.pep	10	20	30	40	50	60
	MNNRHFVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
g008	MNNRHFVIALGSNLDNPAQQIRGALDALSSHPDIRLEQVSSLYMTAPVGYDNQPDFINA					

	10	20	30	40	50	60
m008.pep	70	80	90	100	110	120
	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDDIIDFDGISSDDTRLTLPHPR					
g008	70	80	90	100	110	120
	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLDLDIIDFDGISSDDPRLTLPHPR					
m008.pep	130	140	150	160		
	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX					
	:					
g008	130	140	150	160		
	AHERSFVIRPLAEILPDFILGKYGKVVELSKRLGNQGIRLLPDRX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 55>:

**g009.seq**

```

1   ATGCCCCGCG CTGCCGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GCGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAT TCAGCACACG CAGGCGCGAA ACCAGTCCGT AATGGCGGTA
151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTcg cGTTCCAAGC
201 TGTGTGTTT CAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AaaaGCCATA A

```

This corresponds to the amino acid sequence <SEQ ID 56; ORF 009.ng>:

**g009.pep**

```

1   MPRAAVAFAER HHHKSKEQN THRRADAEIA EGFAVGNQHT QARNQSVMAV
51  QLPLVAFSDK VVVAFAQAVVQ AEIQVFADGG KTWQKP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 57>:

**m009.seq**

```

1   ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAT TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTA
151 CAGCTGCCGC CGGTCGCCTT TTCGGATAAA GTGGTTGTTCG CGTTCCAAGC
201 TGTGTGTTT CAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AAAAGCCATA A

```

This corresponds to the amino acid sequence <SEQ ID 58; ORF 009>:

**m009.pep**

```

1   MPRAAVAFAER HHHKSKEQN THRRADAEIA EGFAVGNQHT QARKQSVMAV
51  QLPPVAFSDK VVVAFAQAVVQ AEIQVFADGG KTWQKP*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 009 shows 97.7% identity over a 86 aa overlap with a predicted ORF (ORF 009.ng) from *N. gonorrhoeae*:

**m009/g009**

	10	20	30	40	50	60
m009.pep	MPRAAVAFAERHHHKSKEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK					
g009	MPRAAVAFAERHHHKSKEQNTHRRADAEIAEGFAVGNQHTQARNQSVMAVQLPLVAFSDK					
	10	20	30	40	50	60
m009.pep	70	80				
	VVVAFAQAVVQAEIQVFADGGKTWQKPX					
g009	70	80				
	VVVAFAQAVVQAEIQVFADGGKTWQKPX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 59>:

```
a009.seq
1   ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTC
151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTCT CGTTCCAAGC
201 TGTTCCTTCAG GCGGAAATAC AGGTTTTTCG TGATGGCGGC AAAACGTGGC
251 AAAAGCCATA A
```

This corresponds to the amino acid sequence <SEQ ID 60; ORF 009.a>:

```
a009.pep
1   MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFVGNQHT QARKQSVMAV
51  QLPLVAFSDK VVVAFAQVLQ AEIQVFADGG KTWQKP*
```

m009/a009 97.7% identity over a 86 aa overlap

```

      10      20      30      40      50      60
m009.pep MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK
a009      MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPLVAFSDK
      10      20      30      40      50      60

      70      80
m009.pep VVVAFAQVAVQAEIQVFADGGKTWQKPX
a009      VVVAFAQVAVQAEIQVFADGGKTWQKPX
      70      80
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 61>:

```
g010.seq
1   ATGGGTTTTTC CTGTTGCGAA GTTTGATGCC GTGATTGTCTG GCGGTGGCGG
51  TGCAGGTTTTC CGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTTGTCT TAAAGTGTTC CCGACCCGCT CGCATACCGT AGCGGCGCAG
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTCTG GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
351 TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTCT TATTTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGGTTTGG GCATTTGCGC CCGTGCGGGC ATTCCGTTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTGAacgc
801 cgacggcgaa cgcTTTATGG AAcgctatgc GCcgACCGta aAgaCTTGG
851 CTTCTCGCga cgtGGTTTCA CgcgcGatgG CGatggaAat ctatgaaggt
901 cgcggtctgTG GtaaaaAcaA agaCCacgtC TTAAGTAAAA TCGACcAtAt
951 cggTGcAGAA AAAATTATGG AAAAATGCC GGGCATCCGC GAGATTTCCA
1001 TTCagtttgc cGGTATCGAT CCGATTAAAG ACCCGATTcc ggttgTGCCG
1051 ACTACCCACT ATATGATGGG CGGCATTcGg aCCAATTATC ACGGTGAAGT
1101 TGTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT
1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga
```

This corresponds to the amino acid sequence <SEQ ID 62; ORF 010.ng>:

g010.pep

```

1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCAVLSKVF PTRSHTVAAQ
51  GGISASLGNV QEDRWDWHMY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPPFDRVE SGKIYORPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNV RANTQF FVEWTAQDLI RDENG DVVG V TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCGKKNKDHV LLKIDHIGAE KIMEKLP GIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEV PVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF RPTPR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 63>:

**m010.seq (PARTIAL)**

```

1  ..nTCCAATTAT CCAAATCCGG TCTGAATTGT GCCGTTTTGT CTAAAGTGTT
51  CCCGACCCGT TCGCATACCG TAGCGGCGCA GGGCGGTATT TCCGCTCTn
101 TGGGTAATGT GCAGGAAGAC CGTTGGGACT GGCACATGTA CGATACCGTG
151 AAAGGTTCCG ACTGGTTGGG CGACCAAGAT GCGATTGAGT TTATGTGCCG
201 CGCCGCGCCT GAAGCCGTAA TTGAGTTGGA ACACATGGGT ATGCCTTTTG
251 ACCGTGTGGA AAGCGGTAAA ATTTATCAGC GTCCTTTCGG CGGCCATACT
301 GCCGAACACG GTAAACGCGC GGTAGAACGC GYCTGTGCGG TTGCCGACCG
351 TACAGGTCAT GCGATGCTGC ATACTTTGTA CCAACAAAAC GTCCGTGCCA
401 ATACGCAATT CTTTGTGGAA TGGACGGCAC AAGATTTGAT TCGTGATGAA
451 AACGGCGATG TCGTCGGCGT AACCGCCATG GAAATGGAAA CCGGCGAAgT
501 TTATATTTTC CACGCTAAAG CTGTGATGTT TGCTACCGGC GCGGCGGTC
551 GTATTTATGC GTCTTCTACC AATGCCTATA TGAATACCGG CGATGGTTTG
601 GGTATTTGTG CGCGTGCAGG TATCCCGTTG GAAGACATGG AATTCTGGCA
651 ATTCAGCCG ACCGGCGTGG CGGGTGCGGG CGTGTTGATT ACCGAA....

```

This corresponds to the amino acid sequence <SEQ ID 64; ORF 010>:

**m010.pep (PARTIAL)**

```

1  ..XQLSKSGLNC AVLSKVF PTR SHTVAAQGGI SASXGNVQED RWDWHMYDTV
51  KGS DWLG DQD AIEFM CRAAP EAVIELEHMG MPFDRVESGK IYQRPFGGHT
101 AEHGKRAVER XCAVADRTGH AMLHTLYQQN VRANTQFFVE WTAQDLIRDE
151 NGDVVGVTAM EMETGEVYIF HAKAVMFATG GGGRIYASST NAYMNTGDGL
201 GICARAGIPL EDM EFWQFQF TGVAGAGVLI TE...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 65>:

**a010.seq**

```

1  ATGGGCTTTC CTGTTGCGAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTTGTC TAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
451 CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GGCGGTATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTA CTGAAAA TCGACCATAT
951 CGGCGCAGAA AAAATTATGG AAAA ACTGCC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051 ACTACCCACT ATATGATGGG CGGTATTCCG ACCAACTACC ATGGCGAAGT
1101 TGTCGTTCC TCAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAAGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAACCT TTGCCTGCTA

```

```

1301 ATGCCGGCGA ACTGACCCGC CAACGTATCG AGCGTTTGGG CAATCAAAC
1351 GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401 ATTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTAATTGAA GTGGCGAAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATGCCAATA CCTTGTCCCTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGA

```

This corresponds to the amino acid sequence <SEQ ID 66; ORF 010.a>:

```

a010.pep
1  MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCVLSKV F PTRSHTVAAQ
51  GGISASLGNV QEDRWDMHY DTVKGSWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRRANTOF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RCGGKNKDHV LLKIDHIGAE KIMEKLPGR EISIQFAGID PIKDPPIVVP
351 TTHYMMGGIP TNYHGEVVPV QGDEYEVVPK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMIAA ERVKRTEIKD
501 KSKVWNTARI EALELDNLE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*

```

m010/a010 98.7% identity over a 231 aa overlap

```

m010.pep
10 20 30
XQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASXGNV
a010
10 20 30 40 50 60
MGFPVRKFDAVIVGGGGAGLRAXLQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV

40 50 60 70 80 90
QEDRWDMHYDTVKGSDWLG DQDAIEFMCR AAPEAVIELE HMGMPFDRVESGKIYQRPFG
a010
70 80 90 100 110 120
QEDRWDMHYDTVKGSDWLG DQDAIEFMCR AAPEAVIELE HMGMPFDRVESGKIYQRPFG

100 110 120 130 140 150
GHAEHGKRAVERXCAVADR TGHAMLHTLY QQNVRRANTOFFVEWTAQDLIRDENG DVVGV
a010
130 140 150 160 170 180
GHAEHGKRAVERACAVADR TGHAMLHTLY QQNVRRANTOFFVEWTAQDLIRDENG DVVGV

160 170 180 190 200 210
TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
a010
190 200 210 220 230 240
TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ

220 230
FHPTGVAGAGVLITE
a010
250 260 270 280 290 300
FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVS RAMAMEIYEG

```

Computer analysis of this amino acid sequence gave the following results:



Homology with a predicted ORF from *N. gonorrhoeae*

ORF 010 shows 98.7% identity over a 231 aa overlap with a predicted ORF (ORF 010.ng) from *N. gonorrhoeae*:

m010.pep/g010.pep

m010.pep				10	20	30	
				XQLSKSGLNCAVL	SKVFP	TRSHTVAAQGGISAS	XGNV
g010	MGFPVRKFD	DAVIVGGGGAGL	RAALQ	LSKSG	LNC	AVLSKVFP	TRSHTVAAQGGISASLGNV
	10	20	30	40	50	60	
m010.pep	40	50	60	70	80	90	
	QEDRWDWHMYD	TVKGS	DLGDD	DAIEF	MCRAA	PEAVIE	LHMGMPFDRVESGKIYQRPFG
g010	QEDRWDWHMYD	TVKGS	DLGDD	DAIEF	MCRAA	PEAVIE	LHMGMPFDRVESGKIYQRPFG
	70	80	90	100	110	120	
m010.pep	100	110	120	130	140	150	
	GHTAEHGKRA	VERXCA	VADRT	GHAML	HTLYQ	QNVRA	NTQFFVEWTAQDLIRDENG
g010	GHTAEHGKRA	VERACAV	ADRTG	HAMLH	TLYQ	QNVRA	NTQFFVEWTAQDLIRDENG
	130	140	150	160	170	180	
m010.pep	160	170	180	190	200	210	
	TAMEMETGE	VYIFH	AKAVM	FATGG	GGRIY	ASSTN	AYMNTGDGLGICARAGI
g010	TAMEMETGE	VYIFH	AKAVM	FATGG	GGRIY	ASSTN	AYMNTGDGLGICARAGI
	190	200	210	220	230	240	
m010.pep	220	230					
	FQPTGVAG	AGVLI	TE				
g010	FHPTGVAG	AGVLI	TEGVR	GGILL	NADGE	RFME	RYAPT
	250	260	270	280	290	300	
	VDLASR	DVVS	RAMA	MEI	YEG		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 67>:

g010-1.seq..

```

1  ATGGGTTTTTCTGTTTCGCAA GTTTGATGCC GTGATTGTGCG GCGGTGGCGG
51  TGCAGGTTTACGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTTTGTC TAAAGTGTTTCCGACCCGCT CGCATACCGT AGCGGCGCAG
151 GGCAGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTGCG GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
351 TCCTTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTGCT TATTTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGGTTTGG GCATTTGCGC CCGTGCGGGC ATTCGGTTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGGTATTTC TGTTGAACGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTGGTTTTCA CGCGCGATGG CGATGGAAAT CTATGAAGGT
901 CGCGGCTGTG GTAAAAACAA AGACCACGTC TACTGAAAA TCGACCATAT
951 CGGTGCAGAA AAAATTATGG AAAAAGTACC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTTGC CGGTATCGAT CCGATTAAAG ACCCGATTCC GGTGTGCCC
1051 ACTACCACT ATATGATGGG CGGCATTCCG ACCAATTATC ACGGTGAAGT
1101 TGTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT

```

1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga

This corresponds to the amino acid sequence <SEQ ID 68; ORF 010-1.ng>:

g010-1.pep

```

1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKVFPTRSHTVAAQ
51  GGISASLGNV QEDRWDWHMY DTVKGSDDLWQ DQDAIEFMCR AAPEAVIELE
101 HMGMPPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE REMERYAPT V KDLSRDVVS RAMAMEIYEG
301 RGCCKNKDHF LLKIDHIGAE KIMEKLPGR EISIQFAGID PIKDFIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEVVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF RPTPR*

```

g010-1 / P10444

sp|P10444|DHSA\_ECOLI SUCCINATE DEHYDROGENASE FLAVOPROTEIN SUBUNIT  
 gnl|PID|d1015210 (D90711) Succinate dehydrogenase, flavoprotein [Escherichia coli] gi|1786942  
 (AE000175) succinate dehydrogenase flavoprotein subunit [Escherichia coli] Length = 588  
 Score = 1073 (495.6 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169  
 Identities = 191/303 (63%), Positives = 238/303 (78%)

```

Query: 1 MGFPVRKFDAVIVXXXXXXXXXXXXXSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV 60
M PVR+FDVAV++ S+SG CA+LSKVFPTRSHTV+AQGGI+ +LGN
Sbjct: 1 MKLPVREFDAVIVIGAGGAGMRAALQISQSGQTCALLSKVFPTRSHTVSAQGGITVALGNT 60

Query: 61 QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG 120
ED W+WHMYDTVKGSD++GDQDAIE+MC+ PEA++ELEHMG+PF R++ G+IYQRPFG
Sbjct: 61 HEDNWEWHMYDTVKGSDYIGDQDAIEYMCKTGPEAILELEHMGPFRLDDGRIYQRPFG 120

Query: 121 GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRRANTQFFVEWTAQDLIRDENGDVVGV 180
G + G R A ADRTGHA+LHTLYQQN++ +T F EW A DL+++++G VVG
Sbjct: 121 GQSKNFGGEQAARTAAADRTGHALLHTLYQQNLKNHTTIFSEWYALDLVKNQDGAUVGC 180

Query: 181 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ 240
TA+ +ETGEV F A+A + ATGG GRIY S+TNA++NTGDG+G+ RAG+P++DME WQ
Sbjct: 181 TALCIETGEVVYFKARATVLTAGGAGRIYQSTTNAHINTGDGVGMAIRAGVPVQDMMEMWQ 240

Query: 241 FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLSRDVVS RAMAMEIYEG 300
FHPTG+AGAGVL+TEG RGEVG LLN GERFMERYAP KDLA RDVV+R++ +EI EG
Sbjct: 241 FHPTGIAGAGVLVTEGCRGEGGYLLNKHGERFMERYAPNAKDLA GRD VVAR SIMIEIREG 300

Query: 301 RGC 303
RGC
Sbjct: 301 RGC 303

```

Score = 249 (115.0 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169  
 Identities = 53/102 (51%), Positives = 62/102 (60%)

```

Query: 309 HVLLKIDHIGAEKIMEKLPGIREISIQFAGXXXXXXXXXXXXTTHYMMGGIPTNYHGEVV 368
H LK+DH+G E + +LPGI E+S FA T HYMMGGIPT G+ +
Sbjct: 310 HAKLKL DHLGKEVLESRLPGILELSRTFAHVDVPVKEPIPVIPVCHYMMGGIPTKVGTGQAL 369

Query: 369 VPQGEYEVVKGLYAAGECACASVHGANRLGTNSLLDLVVF 410
+V V GL+A GE AC SVHGANRLG NSLLDLVVF
Sbjct: 370 TVNEKGEDVVVPGFLFAVGEIACSVHGANRLGNSLLDLVVF 411

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 69>:

m010-1.seq..

```

1  ATGGGTTTTTCTGTTCGCAA GTTTGATGCC GTGATTGTGCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCAGCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTTTGTCTAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GCGGTTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCGGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC

```

```
451 CAACAAAACG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGTCG TATTTATGCG TCTTCTACCA ATGCCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGST ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAATGC
801 CGACGCGCAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
951 CGGCGCAGAA AAAATTATGG AAAAAGTACC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051 ACTACCCACT ATATGATGGG CGGCATTCCG ACCAATTACC ACGGCGAAGT
1101 TGTCGTTCGC CAAGGTGAAG ATTACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAGGTGA GTGCGCTTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACCAACTCCC TGTTGGACTT GGTGGTATTC GGTAAAGCTG CCGCGCACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCTGCTA
1301 ATGCAGGTGA GTTGACCCGC CAACGTATCG AGCGTTTGGG CAACCAAACC
1351 GATGGTGAAA ACGTTGATGC ATTGCGTCGC GAACTGCAAC GCTCTGTACA
1401 ACTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTGATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TCGGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATATCAATA CCTGTCTCTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGATGA
```

This corresponds to the amino acid sequence <SEQ ID 70; ORF 010-1>:

m010-1.pep..

```
1 MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCAVLISKVF PTRSHTVAAQ
51 GGISASLGNV QEDRWDWHMY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRANTQF FVEWTAQDLI RDENG DVVG V TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCGKKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVPV QGEDYEVVPK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR ORIERLDNQT
451 DGENVDALRR ELQSVQLHA GVFRTEILS KGVREVMAlA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DINTLSYKPV HTKPLSVEYI KPAKRVY*
```

m010-1 / g010-1 99.5% identity in 410 aa overlap

```
10 20 30 40 50 60
m010-1.pep MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCAVLISKVFPTRSHTVAAQGGISASLGNV
|||||
g010-1 MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCAVLISKVFPTRSHTVAAQGGISASLGNV
10 20 30 40 50 60

70 80 90 100 110 120
m010-1.pep QEDRWDWHMYDTVKGS DWLG DQDAIEFMCR AAPEAVIELEHMGMPFDRVESGKIIYQRPFG
|||||
g010-1 QEDRWDWHMYDTVKGS DWLG DQDAIEFMCR AAPEAVIELEHMGMPFDRVESGKIIYQRPFG
70 80 90 100 110 120

130 140 150 160 170 180
m010-1.pep GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVG V
|||||
g010-1 GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVG V
130 140 150 160 170 180

190 200 210 220 230 240
m010-1.pep TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
|||||
g010-1 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
```

	190	200	210	220	230	240
m010-1.pep	250	260	270	280	290	300
	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVS RAMAMEIYEG					
g010-1	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVS RAMAMEIYEG					
	250	260	270	280	290	300
m010-1.pep	310	320	330	340	350	360
	RGCGKNKDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPFVVPVTHYMMGGIP					
g010-1	RGCGKNKDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPFVVPVTHYMMGGIP					
	310	320	330	340	350	360
m010-1.pep	370	380	390	400	410	420
	TNYHGEVVVPQGEDYEVVPKGLYAAGECACASVHGANRLGNSLLDLVVFKAAGDSMIK					
g010-1	TNYHGEVVVPQGEDYEVVPKGLYAAGECACASVHGANRLGNSLLDLVVFRTPRX					
	370	380	390	400	410	
m010-1.pep	430	440	450	460	470	480
	FIKEQSDWKPLPANAGELTRQRIERLDNQTDGENVDALRRELQRSVQLHAGVFRTEILS					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 71>:

a010-1.seq..

```

1   ATGGGCTTTC CTGTTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTGTGC TAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GCGGTGATTT CCGCTCTCTT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTATCAGCG
351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
451 CAACAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GGCGGTATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTAAGTAAAA TCGACCATAT
951 CGGCGCAGAA AAAATTATGG AAAAAGTACC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051 ACTACCCACT ATATGATGGG CCGTATTCGG ACCAACTACC ATGGCGAAGT
1101 TGTCGTTTCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
1301 ATGCCGCGCA ACTGACCCGC CAACGTATCG AGCGTTTGGG CAATCAAAC
1351 GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAAGTGAAC GCTCCGTACA
1401 ATTGCACGCC GGCGTGTTC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTAATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCG ATGCGCAATA CCTTGTCTTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGA

```

This corresponds to the amino acid sequence <SEQ ID 72; ORF 010-1.a>:

a010-1.pep..

```

1   MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCAVLSKVF PTRSHTVAAQ
51  GGISASLGNV QEDRWDDHMY DTVKGSDDLW DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHAEHGKRA VERACAVADR TGHAMLHTLY

```

151 QQNVRANTQF FVEWTAQDLI RDENGDVVG V TAMEMETGEV YIFHAKAVMF  
201 ATGGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG  
251 VLITEGVRGE GGILLNADGE RFMERYAPT V KDLASRDVVS RAMAMEIYEG  
301 RGCGKNKDHV LLKIDHIGAE KIMEKLP G IR EISIQFAGID PIKDPIPVVP  
351 TTHYMMGGIP TNYHGEVVVP QGDEYEV P V K GLYAAGECAC ASVHGANRLG  
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT  
451 DGENVDALRR ELQRSVQLHA GVFR T DEILS KGVREVM AIA ERVKRTEIKD  
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE  
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY\*

m010-1 / a010-1 99.3% identity in 587 aa overlap

a010-1.pep	10	20	30	40	50	60
	MGFPVRKFD	AVIVGGGGAGL	RAXLQLSK	SGLNCAVL	SKVPFTR	SHTVAAQGGISASLGNV
a010-1	10	20	30	40	50	60
	MGFPVRKFD	AVIVGGGGAGL	RAALQLSK	SGLNCAVL	SKVPFTR	SHTVAAQGGISASLGNV
a010-1.pep	70	80	90	100	110	120
	QEDRW	DWHMYD	TVKGS	DLGDDQ	DAIEFMC	RAAPEAVIELEHMGMPFDRVESGKIYQRPFG
m010-1	70	80	90	100	110	120
	QEDRW	DWHMYD	TVKGS	DLGDDQ	DAIEFMC	RAAPEAVIELEHMGMPFDRVESGKIYQRPFG
a010-1.pep	130	140	150	160	170	180
	GHTAEH	GKRAVER	ACAVAD	RDTGH	AMLHTLY	QQNVRANTQFFVEWTAQDLIRDENGDVVG
m010-1	130	140	150	160	170	180
	GHTAEH	GKRAVER	ACAVAD	RDTGH	AMLHTLY	QQNVRANTQFFVEWTAQDLIRDENGDVVG
a010-1.pep	190	200	210	220	230	240
	TAMEMET	GEVYIF	HAKAVM	FATGGG	GRIYAS	SSTNAYMNTGDGLGICARAGIPLEDMEFWQ
m010-1	190	200	210	220	230	240
	TAMEMET	GEVYIF	HAKAVM	FATGGG	GRIYAS	SSTNAYMNTGDGLGICARAGIPLEDMEFWQ
a010-1.pep	250	260	270	280	290	300
	FHPTG	VAGAGV	LITEG	VRGEGG	ILLNAD	GERFMERYAPT VKDLASRDVVS RAMAMEIYEG
m010-1	250	260	270	280	290	300
	FHPTG	VAGAGV	LITEG	VRGEGG	ILLNAD	GERFMERYAPT VKDLASRDVVS RAMAMEIYEG
a010-1.pep	310	320	330	340	350	360
	RGCGK	NKDHV	LLKID	HIGAE	KIMEKLP	GIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP
m010-1	310	320	330	340	350	360
	RGCGK	NKDHV	LLKID	HIGAE	KIMEKLP	GIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP
a010-1.pep	370	380	390	400	410	420
	TNYHGE	VVV	PQGDE	YEV	PKGLYA	AAGECACASVHGANRLGTNSLLDLVVF GKAAGDSMIK
m010-1	370	380	390	400	410	420
	TNYHGE	VVV	PQGED	YEV	PKGLYA	AAGECACASVHGANRLGTNSLLDLVVF GKAAGDSMIK
a010-1.pep	430	440	450	460	470	480
	FIKEQ	SDWKPL	PANAGEL	TRQRI	ERLDN	QTDGENVDALRR ELQRSVQLHAGVFR TDEILS
m010-1	430	440	450	460	470	480
	FIKEQ	SDWKPL	PANAGEL	TRQRI	ERLDN	QTDGENVDALRR ELQRSVQLHAGVFR TDEILS
a010-1.pep	490	500	510	520	530	540
	KGVR	EVMAIA	ERVKR	TEIKD	KSKVW	NTARIEALELDNLIE VAKATLVSAE ARKESRGAHA
m010-1	490	500	510	520	530	540
	KGVR	EVMAIA	ERVKR	TEIKD	KSKVW	NTARIEALELDNLIE VAKATLVSAE ARKESRGAHA
a010-1.pep	550	560	570	580		
	SDDH	PERDDE	NWMKHT	LYHSD	ANTLSY	KPVHTKPLSVEYIKPAKRVYX

m010-1

SDDHPERDDENWMKHTLYHSDINTLSYKPVHTKPLSVEYIKPAKRVYX  
 550 560 570 580

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 73>:

g011.seq  
 1 ATGAAGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC  
 51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA  
 101 GCCTGAAAAC CCGCCTTACC GAAGATATGA AAACCGCGAT GCGCGCCAAA  
 151 GATCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAATGCCG CCGTCAAACA  
 201 GTTTGAAGTA GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA  
 251 TCCTGACCAA AATGGTCAA CAGCGCAAAG ACGGCGCGAA AATCTACACT  
 301 GAAGCCGGCC GTCAGGATT GGCAGACAAA GAAAACGCCG AAATCGACGT  
 351 GCTGCACCGC TACCTGCCGC AAATGCTCTC CGCCGGCGAA ATCCGCACCG  
 401 CCGTCGAAGC AGCCGTTGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG  
 451 GGCAAAGTGA TGGTCGTATT GAAAaccGC CTCGCCGCA AAGccgATAT  
 501 GGGCGAAGTC AACAAATCT TGAAAaccGt aCTGACCGCC tga

This corresponds to the amino acid sequence <SEQ ID 74; ORF 011.ng>:

g011.pep  
 1 MKTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKTRLT EDMKTAMRAK  
 51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDGAKIYT  
 101 EAGRQDLADK ENAEIDVLHR YLPQMLSAGE IRTAVEAAVA ETGAAGMADM  
 151 GKVMVVLKTR LAGKADMGEV NKILKTVLTA \*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 75>:

m011.seq (partial)  
 1 ATGAGGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC  
 51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA  
 101 GCCTGAAAAT CCGCCTTACC GAAGACATGA AAACCGCGAT GCGCGCCAAA  
 151 GACCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAACGCCG CCGTCAAACA  
 201 GTTTGAAGTG GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA  
 251 TCCTGACCAA AATGGTCAA CAGCGAAAAG ACAGCGCGAA AATCTACACT  
 301 GAAGCCGGCC GTCAGGATT GGCAGACAAA GAAAACGCCG AAATCGAGGT  
 351 ACTGCACCGC TACCTTCCCC AAATGCTTTC CGCCGGCGAA ATCCGTACCG  
 401 AGGTCAAGC TGCCGTTGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG  
 451 GGTAAAGTCA TGGGGCTGCT GAAAACCGCG CTCGCAGGTA AAGCCGA...

This corresponds to the amino acid sequence <SEQ ID 76; ORF 011>:

m011.pep (partial)  
 1 MRTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKIRLT EDMKTAMRAK  
 51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDSAKIYT  
 101 EAGRQDLADK ENAEIEVLHR YLPQMLSAGE IRTEVEAAVA ETGAAGMADM  
 151 GKVMGLLKTR LAGKA.....

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 011 shows 95.8% identity over a 165 aa overlap with a predicted ORF (ORF 011.ng) from *N. gonorrhoeae*:

m011/g011

	10	20	30	40	50	60
m011.pep	MRTHRKTCSAVCF	AFQTASKPAVS	IRHPSEDI	MSLKIRLT	EDMKTAMRAK	DQVSLGTIRL
	:					
g011	MKTHRKTCSAVCF	AFQTASKPAVS	IRHPSEDI	MSLKTRLT	EDMKTAMRAK	DQVSLGTIRL
	10	20	30	40	50	60
	70	80	90	100	110	120
m011.pep	INAAVKQFEVDER	TEADDAKITAI	LTKMVKQRK	DSAKIYTEAG	RQDLADKEN	AEIEVLHR
g011	INAAVKQFEVDER	TEADDAKITAI	LTKMVKQRK	DGAKIYTEAG	RQDLADKEN	AEIDVLHR

	70	80	90	100	110	120
	130	140	150	160		
m011.pep	YLPQMLSAGEIRTEVEAAVAETGAAGMADMGKVMGLLKTRLAGKA					
g011	YLPQMLSAGEIRTAVEAAVAETGAAGMADMGKVMVVLKTRLAGKADMGKVNKILKTVLTA					
	130	140	150	160	170	180
g011	X					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 77>:

**g012.seq**

```

1   ATGCTCGCCC GTCGCTATTT TTTCAATATC CAACCCGGGG CGGTTTTTCAC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGCCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACa
201 gGcggTGGAT ATTCGgcact tccgCcacca caccacccga accgatgacc
251 gcaaacggaG CGGAACAAT TTTATCCGCC acacacgccca tcatatagcc
301 gcCGCTTGCC GCGACCTTAT CGAcggcgac ggTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551 GCAGATTTCT CCCCGCCCTC CTTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CTTTTTTTTC CTGATGTTTT GTCTCTTCCT
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 78; ORF 012.ng>:

**g012.pep**

```

1   MLARRYFFNI QPGAVFTDKL LEQLMRFLQF LPEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRHFRHHTHR TDDRKRSGNN FIRHTRHHIA
101 AACRDLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS EQNLIFRLGN
151 QKHRRLMTQ GFYGVCIQIA VKIQHKKAGF LRFRGRFLPAL LQTLFLCFGF
201 RLFLFLFFFF LMFCLFLA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 79>:

**m012.seq**

```

1   ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACa
201 GGCGGTGGAT ATTCGGTACT TCCGCCACCA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 GCCGCTCGCn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
351 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
401 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
451 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
501 nnnnnnnnnn nnnnnnnnnC AACACAAAAA GGCGTGATTT nTGCGTTTCG
551 GCAGATTTCT CCCACCCCTC CTTCAAACGT TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTGT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 80; ORF 012>:

**m012.pep**

```

1   MLARCHFLNI QLR AVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRYFRHHTHR TDNRKRSGSN FIRHTRHHIT
101 AARXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
151 XXXXXXXXXXX XXXXXXXXXXX XXXQHKKAF XRFRGRFLPTL LQTFFLCFGF
201 RLFLFLFLFF LMLCLFPA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 81>:

```
a012.seq
1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTGCGGAAT
101 TTCTGTTTGC CTTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GCGGTGGAT ATTCCGTACT TCCGCTACAA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCGGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTACGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551 GAAGATTTCT CCCCACCCTC CTTCAAACGC TTTTCTCTCT CTTTGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTTT GCCTCTTCCC
651 CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 82; ORF 012.a>:

```
a012.pep
1  MLARCHFLNI QLRVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRYFRYNTHR TDNRKRSNN FIRHTRHHIT
101 TARRHLIDGD GQRNIAFAQT PKLRSQTVT VNHAARTFQS KQNLIFRLGN
151 QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFRFLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFLCLFPA*
```

m012/a012 64.2% identity over a 218 aa overlap

m012.pep	10	20	30	40	50	60
	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
a012						
	10	20	30	40	50	60
	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
m012.pep	70	80	90	100	110	120
	NIMFFQQAVDIRYFRHHTHRTDNRKRSNFIHTRHHITAARXXXXXXXXXXXXXXXXXXXX					
a012	:     :     :     :     :					
	70	80	90	100	110	120
	NIMFFQQAVDIRYFRYNTHRTDNRKRSNFIHTRHHITTARRHLIDGDGQRNIAFAQT					
m012.pep	130	140	150	160	170	180
	XXXQHKKAXF					
a012	: : : : :					
	130	140	150	160	170	180
	PKLRSQTVTVNHAARTFQSKQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF					
m012.pep	190	200	210	219		
	XRFRFLPTLLQTFFLCFGRFLFLFLFLFLMLCLFPAX					
a012	:     :     :     :					
	190	200	210			
	LRFRFLPTLLQTLFLCFGRFLFLFLFLFLMLCLFPAX					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 012 shows 58.7% identity over a 218 aa overlap with a predicted ORF (ORF 012.ng) from *N. gonorrhoeae*:

m012/g012

m012.pep	10	20	30	40	50	60
	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
g012	:     :     :     :					
	MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI					



195

	10	20	30	40	50	60
m012.pep	70	80	90	100	110	120
	NIMFFQQAVDIRYFRHHHTHRTDNRKRSGSNFIRHTRHHITAARXXXXXXXXXXXXXXXXXXXX					
g012	NIMFFQQAVDIRHFRHHHTHRTDDRKRSGNNFIRHTRHHIAAACRDLDGDGQRNIAFAQT					
	70	80	90	100	110	120
m012.pep	130	140	150	160	170	180
	XXQHKKAXF					
g012	PKLRSRQTVTVNHAARTFQSEQNLIIFRLGNQKHRRNLMTQGFGYVCIQIAVKIQHKKAGF					
	130	140	150	160	170	180
m012.pep	190	200	210	219		
	XRFGRLPTLLQTFFLCFGRFLFLFLFLFFLMLCLFPAX					
g012	LRFGRLPALLQTLFLCFGRFLFLFLFFFLMFLCLFLAX					
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 83>:

m012-1.seq

1	ATGCTCGCCC	GTTGCCACTT	CCTCAATATC	CAATTGAGGG	CGGTTCTCGC
51	TGACAACTG	CTTGAACAAC	TGATGCGTTT	CCTCCAGTTC	CTGTCGGAAT
101	TTCTGTTTGC	CCTTTTCCGT	ATTTTCACCC	ATAAAAGTAA	CCGTGCGCTT
151	AAATTCGCCC	GCCGTCATCA	CATCCACATC	AATATCATGT	TTTTTCAACA
201	GGCGGTGGAT	ATTCGGTACT	TCCGCCACCA	CACCCACCGA	ACCGACAATC
251	GCAAACGGAG	CGGAAGCAAT	TTTATCCGCC	ACACACGCCA	TCATATAAACC
301	GCCGCTCGCC	GCCACCTTAT	CGACGCGGAC	GGTCAGCGGA	ATATTGCGTT
351	CGCGCAAACG	CyTAAGCTGC	GAAGCCGCCA	AACCGTAACC	GTGAACCACG
401	CCGCCCCGAC	TTTCCAATCT	GAGCAGAACC	TCATCTTCAG	GCTTGGCAAT
451	CAAAAGCACC	GCCGTAATCT	CATGACGCAA	GGATTCTACG	GCGTGTGCAT
501	ACAAATCGCC	GTCAAAATCC	AACACAAAAA	GGCGGGATT	TTGCGTTTCG
551	GCAGATTCT	CCCCACCCTC	CTTCAAACGC	TTTTTCTCTG	CTTTGGCTTC
601	CGCCTTTCC	TTTTTCTTTT	CCTCTTTTTC	CTGATGTTTT	GCCTCTTCCC
651	CGCTTAA				

This corresponds to the amino acid sequence <SEQ ID 84; ORF 012-1>:

m012-1.pep

1	MLARCHFLNI	QLRAVLADKL	LEQLMRFLQF	LSEFLFALFR	IFTHKSNRAL
51	KFARRHHIHI	NIMFFQQAVD	IRYFRHHHTH	TDNRKRSGSN	FIRHTRHHIT
101	AARRHLIDGD	GQRNIAFAQT	XKLRSRQTVT	VNHAARTFQS	EQNLIIFRLGN
151	QKHRRNLMTQ	GFGYVCIQIA	VKIQHKKAGF	LRFGRLPTL	LQTLFLCFGF
201	RLFLFLFLFF	LMFLCLFPA*			

m012-1/g012 91.7% identity in 218 aa overlap

	10	20	30	40	50	60
m012-1.pep	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
g012	MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI					
	10	20	30	40	50	60
m012-1.pep	70	80	90	100	110	120
	NIMFFQQAVDIRYFRHHHTHRTDNRKRSGSNFIRHTRHHITAARRHLIDGDGQRNIAFAQT					
g012	NIMFFQQAVDIRHFRHHHTHRTDDRKRSGNNFIRHTRHHIAAACRDLDGDGQRNIAFAQT					
	70	80	90	100	110	120
m012-1.pep	130	140	150	160	170	180
	XKLRSRQTVTVNHAARTFQSEQNLIIFRLGNQKHRRNLMTQGFGYVCIQIAVKIQHKKAGF					
g012	PKLRSRQTVTVNHAARTFQSEQNLIIFRLGNQKHRRNLMTQGFGYVCIQIAVKIQHKKAGF					

	130	140	150	160	170	180
	190	200	210	219		
m012-1.pep	LRFGRFLPTLLQTLFLCFGRFLFLFLFFFLMFCLFPAX					
	:     :     :     :					
g012	LRFGRFLPALLOTLFLCFGRFLFLFLFFFLMFCLFLAX					
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 85>:

a012-1.seq

```

1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCACAA
201 GGCGGTGGAT ATTGCTACT TCCGCTACAA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGCGGAC GGTACAGCGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCCTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGCCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551 GAAGATTTC CCCCACCTC CTTCAAACGC TTTTCTCTG CTTGGCTTC
601 CGCCTTTTC TTTTCTTTT CCTCTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 86; ORF 012-1.a>:

a012-1.pep

```

1  MLARCHFLNI QLRVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRYFRYNTHR TDNRKRSNN FIRHTRHHIT
101 TARRHLIDGD GQRNIAFAQT PKLSRQTVT VNHAARTFQS KQNLIFRLGN
151 QKRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFCLFPA*

```

a012-1/m012-1 97.2% identity in 218 aa overlap

	10	20	30	40	50	60
a012-1.pep	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
m012-1	:     :     :     :					
	10	20	30	40	50	60
a012-1.pep	NIMFFQQA VDIRYFRYNTHRTDNRKRSNNFIRHTRHHITARRHLIDGDGQRNIAFAQT					
m012-1	:     :     :     :					
	70	80	90	100	110	120
a012-1.pep	NIMFFQQA VDIRYFRYNTHRTDNRKRSNNFIRHTRHHITARRHLIDGDGQRNIAFAQT					
m012-1	:     :     :     :					
	70	80	90	100	110	120
a012-1.pep	PKLSRQTVTVNHAARTFQSKQNLIFRLGNQKRRNLMTQGFYGVCIQIAVKIQHKKAGF					
m012-1	:     :     :     :					
	130	140	150	160	170	180
a012-1.pep	PKLSRQTVTVNHAARTFQSKQNLIFRLGNQKRRNLMTQGFYGVCIQIAVKIQHKKAGF					
m012-1	:     :     :     :					
	130	140	150	160	170	180
a012-1.pep	LRFGRFLPTLLQTLFLCFGRFLFLFLFFFLMFCLFPAX					
m012-1	:     :     :     :					
	190	200	210	219		
a012-1.pep	LRFGRFLPTLLQTLFLCFGRFLFLFLFFFLMFCLFPAX					
m012-1	:     :     :     :					
	190	200	210			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 87>:

g013.seq

```

1  aTgcctttga ccatgctgtg cagcaGGAcg tGCGGTTtgt tcataatata
51  gtCcgaccGG AAAagcggAG GAAaCGCAGT GCCGCGCCCT TCCCTTTCT
101 TGCCGTGGCA GGCGATGCag tTgGATTTCGT AACTTTTTTG CCCTTTtGtc

```

g013.pcp

```

1  MPLTMLCSRT  CGLFIIQSDR  KSGGNAVPRP  SPFLPWQAMQ  LDSYTFCPFV
51  MMLLSAAEAA  AQRQHMKMAV  GSRVVFIVGS  PNVLKPCFLI  LPLRGEKFFW
101 PKSGI*

```

m013.seq

1	ATGCCTTTGA	CCATGCTGTG	ÇAGCAGCACC	TGCGGTTTTT	TCATGATGAA
51	GTCGGAGCGG	TAGAGCGGCG	GAAACATGGT	TCCGCGGGTT	TCGCCCTTTT
101	TGCCGTAGCG	GGCGACGCAG	TTGGATTTCG	ACACTTTTCT	CCCTTTTGTG
151	ATGATGCTGT	TGTCGGCGCG	GGAAGCGGCG	CGCGAGAAGC	AGCCCAAGAC
201	GAGGGCGGTC	GGCAGTCGGG	TTGTGTTTCAT	TGGTGTTTCC	TTCATGTTTG
251	AAACCTTGTT	GTTGATTTTG	CGTAGCGGGT	GAAAGATTTT	TTTGCCGAAT
301	CAGTAG				

m013.ppt

1 MPLTMLCSST CGFFMMKSER XSGGNMVRP SPFLPWQATQ LDSYTFCPFV  
51 MMLLSAAEAA AQQQPKTRAV GSRVVFIVGS FMFETLLIL RSGXKIFLPN  
101 Q\*

a013.seq

1	ATGCCCTTTGA	CCATGCTGTG	CAGCAGCACC	TGCGGTTTTT	TCATGATGAA
51	GTCGGAGCGG	TAGAGCGGCG	GAACATGGT	TCCGCGGCCT	TCGCCCTTTT
101	TGCCGTGGCA	GGCAGCGCAG	TTGGATTCGT	ACACTTTTGT	CCCTTTTGTG
151	ATGATCTGTG	TGTGCGCGCG	AGAAGCCGCG	CGCGAGAGGC	AGCCCAAGAC
201	GAGGGCGGTC	GGCAGTCGGG	TTGTGTTTCAT	TGGTGTTCCT	TTAATGTTTG
251	AAACCTTGTT	GTTGATTTTG	CGTAGCGGGT	GAAGATTTT	CTGCCGAAT
301	CGGTAG				

a013.pep

1 MPLTMLCSST CGFFMMKSER \*SGGNMVERP SPFLPWQATQ LDSYTFCPFV  
51 MMLLSAAEAA AQRQPKTRAV GSRVVF<sup>1</sup>IGVS LMFETLL<sup>1</sup>LIL RSG\*KIFLPN  
101 R\*

	10	20	30	40	50	60
m013.pep	MPLTMLCSSTCGFFMMKSERXSGGNMVRPSPFLPWQATQLDSYTFPCPFVMMLLSAAEEA					
a013	MPLTMLCSSTCGFFMMKSERXSGGNMVRPSPFLPWQATQLDSYTFPCPFVMMLLSAAEEA					
	10	20	30	40	50	60
	70	80	90	100		
m013.pep	AQKQPKTRAVGSRVVFIGVSFMFETLLILLRSGXKIFLPNQX					
	:                         :					
a013	AQRQPKTRAVGSRVVFIGVSLMFETLLILLRSGXKIFLPNRX					
	70	80	90	100		

ORF 013 shows 73.3% identity over a 101 aa overlap with a predicted ORF (ORF 013.ng) from *N. gonorrhoeae*:

## m013/g013

	10	20	30	40	50	60
m013.pep	MPLTMLCSSTCGFFMMKSERXSGGNMVPSPFLPWQATQLDSYTFPCPFVMMLLSAAEAA					
g013	MPLTMLCSRTCGLFIIQSDRKSGGNAVPRSPFLPWQAMQLDSYTFPCPFVMMLLSAAEAA					
	10	20	30	40	50	60
	70	80	90	100		
m013.pep	AQKQPKTRAVGSRVVFIVSF-MFETLLILR-SGXKIFLPNQX					
g013	AQRQHKMKAVGSRVVFIVSPNVLPKPCFLILPLRGEKFFWPKSGIX					
	70	80	90	100		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 93>:

g015.seq

1	ATGCAGTATC	TGATTGTCAA	ATACAGCCAT	CAAATCTTCG	TTACCATCAC
51	CATTTTGGTA	TTCAACATCC	GTTTTTCCT	ACTTTGGAAA	AATCCAGAAA
101	AGCCCTTGGT	CGGCTTTTGG	AAAGCACTGC	CCCACCTCAA	CGACACGATG
151	CTGCTGTTTA	CGGGATTGTG	GCTGATGAAG	ATTACCCATT	TCTCCCCGTT
201	CAACGCGCCT	TGGCTCGGCA	CAAAAATCCT	GCTCCTGTTC	GCCTACATCG
251	CACTGGGCAT	GGTAATGATG	CGCGCCCGTC	CGCGTTTCGAC	CAAGTTCTAC
301	ACCGTTTACC	TGCTCGCTAT	GTGTTGCATC	GCCTGCATCG	TTACCTTGC
351	CAAAACCAAA	GTCCTGCCAT	TCTGA		

This corresponds to the amino acid sequence <SEQ ID 94; ORF 015.ng>:

g015.pep

1	MOYLIVKYSH	<u>QIFVTITILV</u>	<u>FNIRFFLLWK</u>	NPEKPLVGFW	KALPHLNDTM
51	LLFTGLWLMK	ITHFSPFNAP	WLGTKILLLE	AYIALGMVMM	RARPRSTKFY
101	<u>TVYLLAMCCI</u>	<u>ACIVYLAKTK</u>	VLPF*		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 95>:

m015.seq (partial)

1	..AAAATCAGAA	AAGCCTTGGC	GGGCTTTTGG	AAGGCACTGC	CCCACCTTAA
51	CGACACCATG	CTGCTGTTTA	CGGGATTGTG	GCTGATGAAA	ATTACCCATT
101	TCTCCCCGTT	CAACGCGCCT	TGGCTCGGTA	CAAAAATCCT	GCTTCTGCTC
151	GCCTATATCG	CATTGGGTAT	GATGATGATG	CGCGCCCGTC	CGCGTTTCGAC
201	CAAGTTCTAC	ACCGTTTACC	TGCTCGCCAT	GTGTTGCGTC	GCCTGCATCG
251	TTACCTTGC	CAAAACCAAA	GTCCTGCCTT	TCTGA	

This corresponds to the amino acid sequence <SEQ ID 96; ORF 015>:

m015.pep (partial)

1	..KIRKALAGFW	KALPHLNDTM	LLFTGLWLMK	ITHFSPFNAP	WLGTKILLLE
51	<u>AYIALGMMM</u>	<u>RARPRSTKFY</u>	<u>TVYLLAMCCV</u>	<u>ACIVYLAKTK</u>	VLPF*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 97>:

a015.seq

1	ATGCAGTATC	TGATTGTCAA	ATACAGCCAT	CAAATCTTCG	TTACCATCAC
51	CATTTTGGTA	TTCAACATCC	GTGTTTTCNT	ACTTTGGAAA	AATCCAGAAA
101	AGCCCTTGGC	GGGCTTTTGG	AAGGCACTGC	CCCACCTTAA	CGACACCATG
151	CTGCTGTTTA	CGGGATTGTG	GCTGATGAAA	ATTACCCATT	TCTCCCCGTT
201	CAACGCGCCT	TGGCTCGGTA	CAAAAATCCT	GCTTCTGCTC	GCCTATATCG
251	CATTGGGTAT	GATGATGATG	CGCGCCCGTC	CGCGTTTCGAC	CAAGTTCTAC
301	ACCGTTTACC	TGCTCGCCAT	GTGTTGCCTC	ACCTGCATCG	TTACCTTGC
351	CAAAACCAAA	GTCCTGCCTT	TCTGA		

This corresponds to the amino acid sequence <SEQ ID 98; ORF 015.a>:

a015.pep

```

1  MOYLIVKYSH QIFVTITILV FNIRVFXLWK NPEKPLAGFW KALPHLNDTM
51  LLFTGLWLMK ITHFSPFNAP WLGTKILLLL AYIALGMMM RARPRSTKFY
101 TVYLLAMCCL TCIVYLAKTK VLPF*

```

m015/a015 96.7% identity over a 91 aa overlap

```

                                10      20      30
m015.pep                      KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                | | | | | | | | | | | | | | | | | |
a015      LIVKYSHQIFVTITILVFNIRVFXLWKNPEKPLAGFWKALPHLNDTMLLFTGLWLMKITH
              10      20      30      40      50      60

              40      50      60      70      80      90
m015.pep    FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFYTVYLLAMCCVACIVYLAKTKVLP
              | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a015      FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFYTVYLLAMCCLTCIVYLAKTKVLP
              70      80      90      100     110     120

m015.pep    FX
              ||
a015      FX

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 015 shows 94.5% identity over a 91 aa overlap with a predicted ORF (ORF 015.ng) from *N. gonorrhoeae*:

m015/g015

```

                                10      20      30
m015.pep                      KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                | | | | | | | | | | | | | | | | | |
g015      LIVKYSHQIFVTITILVFNIRFFLLWKNPEKPLVGFWKALPHLNDTMLLFTGLWLMKITH
              10      20      30      40      50      60

              40      50      60      70      80      90
m015.pep    FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFYTVYLLAMCCVACIVYLAKTKVLP
              | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g015      FSPFNAPWLGTKILLLLFAYIALGMVMMRRARPRSTKFYTVYLLAMCCVACIVYLAKTKVLP
              70      80      90      100     110     120

m015.pep    FX
              ||
g015      FX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 99>:

g018.seq

```

1  atGCAGCAGG GGCagttggt tggacgcgtc gcccgcaata AAGATATGCG
51  GAATgctggt CTGCATggtC AGCGGATCGG CAACGGGtac gccgcgcgcg
101 tctttgTCGA TATTGATGTT TTCAAACCG ATATgtTCAA CGTTCGGACG
151 GCgACCTACG GCTGCCAACA TATATTGGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT tctACATTGC CGTCTGCGTC GAGTTTGACC
251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT tctTCTCCGA ACACGGCTTT
301 CGCCTCGTCT GA

```

This corresponds to the amino acid sequence <SEQ ID 100; ORF 018.ng>:

g018.pep

```

1  MQQGQLVGRV ARNKDMRNAG LHGQRIGNGY AARVEVDIDV FQTDIVNVRT
51  ATYGCQHIFG NKYAFAILL PMDFYIAVCV EFDLGFSIQM QFQFFSEHGF

```

101 RLV\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 101>:

```
m018.seq
1 ATGCAGCAGA GGCAGTTGGT TGGACGCATC GCCTGCGATG AAGATATGCG
51 GAATACTGGT CTGCATGGTC AGCGGGTCGG CAACAGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATATT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCCCCACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCATC GAGTTTGACC
251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT TCTTCGCCGA ACACGGCGTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 102; ORF 018>:

```
m018.pep
1 MQQRQLVGRI ACDEDMRNTG LHGQRVGNRY AARIFFDIDI FQTDIVNVRT
51 AAHGCGHIFG NKYAFAILL PMDFYIAVCI EFDLGFSIQM QFQFFAEHGV
101 RLV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 103>:

```
a018.seq
1 ATGCAGCAGG GGCAGTTGGT TGGACGCGTC GCCCGCAATA AAGATATGCG
51 GAATACTGGT CTGCATAGTC AGCGGATCGG CAACGGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATGTT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCTACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCGTC GAGTTTGGCC
251 TCGGTTTTAG CATCCAAATG CAGTTTCAAT TCTTCACCGA ACACGGCTTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 104; ORF 018.a>:

```
a018.pep
1 MQQGQLVGRV ARNKDMRNTG LHSQRIGNGY AARIFFDIDV FQTDIVNVRT
51 AAYGCGHIFG NKYAFAILL PMDFYIAVCV EFGLGFSIQM QFQFFTEHGF
101 RLV*
```

m018/a018 86.4% identity over a 103 aa overlap

	10	20	30	40	50	60
m018.pep	MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCGHIFG					
a018	MQQGQLVGRVARNKDMRNTGLHSQRIGNGYAARIFFDIDVFQTDIVNVRTAAYGCGHIFG					
	10	20	30	40	50	60
	70	80	90	100		
m018.pep	NKYAFAILLPMDFYIAVCI EFDLGFSIQM QFQFFAEHGVRLVX					
a018	NKYAFAILLPMDFYIAVCV EFGLGFSIQM QFQFFTEHGFRLVX					
	70	80	90	100		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 018 shows 84.5% identity over a 103 aa overlap with a predicted ORF (ORF 018.ng) from *N. gonorrhoeae*:

m018/g018

	10	20	30	40	50	60
m018.pep	MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCGHIFG					
g018	MQQGQLVGRVARNKDMRNAGLHGQRVGNRYAARVFDIDVFQTDIVNVRTATYGCQHIFG					
	10	20	30	40	50	60

	70	80	90	100
m018.pep	NKYAFFAILLPMDFYIAVCI EFDLGF SIQM QF FFAEHGVRLVX			
	:     :     :			
g018	NKYAFFAILLPMDFYIAVCEFDLGF SIQM QF FSEHGFRVLX			
	70	80	90	100

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 105>:

g019.seq (partial)

```

1  ..ctgctggcgg ccctggtgct tgccgcgtgt tcttcgACAA ACacactTGCC
51  AGCCGCAAG ACCCCGGCAG ACAATATAGA AActgcCgAC CTTCGGCAA
101 GCGTCCAC ccgcCCTGCC GAACCGGAAG GAAAAACGCT GGCAGATTAC
151 GCGGCTACC CGTCCGCACT GGATGCAGTG AAACAGAACA ACGATGCGGC
201 AGCCGCCGCC TATTTGGAAA Acgcaggaga cagCGcgatg gcGGAAAatg
251 tccgaagga gtgGCTGa

```

This corresponds to the amino acid sequence <SEQ ID 106; ORF 019.ng>:

g019.pep (partial)

```

1  ..LLAALVLAAC SSTNTLPAGK TPA DNIETAD LSASVPTRPA EPEGKTLADY
51  GGYPSALDAV KQNDAAAAA YLENAGDSAM AENVRKEWL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 107>:

m019.seq (partial)

```

1  ATGTACCTAC CCTCTATGAA GCATTCCCTG CCGCTGCTGG CGGCCCTGGT
51  GCTTGCCGCG TGTTCTTCGA CAAACACACT GCCAGCCGCG AAGACCCCGG
101 CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCCGCCCT
151 GCCGAACCCG AAAGAAAAAC GCTGGCAGAT TACGGCGGCT ACCCGTCCGC
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
301 AAGTCTTGG GCGCACGCG ACAGTGGACG CTGTTTGAC AGGAATACGC
351 CAAACTCGAA CCGGCAGGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCGCTG AACTGGTCAA AAATACGGGC
451 AACTGCCTT CGGGCTGCAC CAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
601 TTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAAGAA GCACGCAAT GCACGCAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCT TGGACTATTA
801 CGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGCCTTGCGC GCCCGACGTT GGGACGAGCT GGCCTCCGTT
901 ATCTCGCATA TGCCCGAAAA ACTGCAAAAA AGCCCGACCT GGCTCTACTG
951 GCTGGCACGC AGCCGCGCCG CAACGGGCAA CACGCAAGAG GCGGAAAAAC
1001 TTTACAAACA GGCGGCAGCG ACGGGCAGGA ATTTTATATG GGTGCTGGCA
1051 GGGGAAGAAT TGGGTCGGAA AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAAAC AGCGTCCGCC GCATGGCGGA AGACGGTGCA GTCAAACGCG
1151 CACTGGTACT GTTCCAAAAC AGCCAATCTG CCGGTGATGC AAAAATGCGC
1201 CGTCAGGCTC AGGCGGAATG GCGTTTGGC ACACGCGGCT TTGACGAAGA
1251 CAAGCTGCTG ACCGCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCAGC GCAAACCTAA CTACACCTTG
1351 CGCTATATTT CGCCGTTTAA AGACACGGTA ATCCGCCACG CGCAAAATGT
1401 TAATGTCGAT CCGGCTTGGG TTTATGGGCT GATTCGTCAG GAAAGCCGCT
1451 TCGTTATAGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
1501 ATGCTGCCA CGCGCGCGA AATCGCCGCG AAAATCGGTA TGGATGCCCG
1551 ACAACTTTAC ACCGCCGACG GG...

```

This corresponds to the amino acid sequence <SEQ ID 108; ORF 019>:

m019.pep (partial)

```

1  MYLPSMKHSL PLLAALVLAA CSSTNTLPAG KTPADNIETA DLSASVPTRP

```

```

51  AEPERKTLAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARRQWT LFAQEYAKLE PAGRAQEVEC YADSSRNDYT RAAELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWGV
251 GHYQSONLNV PAALDYYGKV ADRRLTDDQ IEWYARAALR ARRWDELASV
301 ISHMPEKLQK SPTWLYWLR SRAATGNTQE AEKLYKQAAA TGRNFYAVLA
351 GEELGRKIDT RNNVPDAGKN SVRRMAEDGA VKRALVLFQV SOSAGDAKMR
401 RQAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNYYL
451 RYISPFKDTV IRHAQNVNVD PAWVYGLIRQ ESRFVIGAQS RVGAQGLMQV
501 MPATAREIAG KIGMDAAQLY TADG...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 109>:

```

a019.seq
1  ATGTACCCAC CCTCTCTGAA GCATTCCCTG CCGCTGCTGG TGGNCCTGGT
51  GCTTGCCGCG TGTCTTNGA CAAACACACT GTCAGCCGAC AAGACCCCGG
101 CAGACAATAT AGAACTGCC GACCTTTCGG CAAGCGTTCC CACCNGCCCT
151 GCCGAACCCG AANGAAAAAC GTNGGCAGAT TACGGCGGCT ACCCGTCCGC
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTGG
251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
301 AAGTCTTTGG GCGCGCGCAG ACAGTGGACG CTGTNTGCAC ANGAATATGC
351 NAAACTCGAA CCGGCANGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCGCCG AACTGGTCAA AAATACGGGC
451 AAAGTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
551 GCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGATT GGCAGCCCG
601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCT TGGACTATTA
801 NGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGCNNNTNNGC NNNCGNNGTT NGNANGANNT GGCNNCCGNN
901 ANCNCGNNNN TGCNNGANAA ACNNNNNNAN AGNCNNANNT NGNTNNANTG
951 NNTGGCACGC AGCCGCGCCG CNACGGGCAA CACGCAANAN CGCGANAAAC
1001 TNTACAAACA GCGGCGAGCA NCGGGCANGA ATTTTATGC NGTGCTGNCN
1051 GGGGAAGAGT TGGGGCGCAN AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAANC AGCGTCTCC GTATGGCGGA AGACGGCGCG ATTAAGCGCG
1151 CGCTGGTGCT GTTCCGAAAC AGCCGAACCG CCGGCGATGC GAAAATGCGC
1201 CGTCNGGCTC AGGCGGAATG GCGTTTCGCC ACACGCGGCT TCGATGAAGA
1251 CAAGCTGCTG ACCGCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACTCAA CTACACCTG
1351 CGCTACATTT CGNNNNNTNA NGACACGGTA ATCCGCCACG CGCAAAATGT
1401 TAATGTCGAT CCGGCGTGGG TTTACGGGCT GATTGTCAG GAAAGCCGCT
1451 TCGTTATGGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
1501 ATGCTGCCCA CCGCGCGCGA AATCGCCGCG AAAATCGGTA TGGATGCCGC
1551 ACAACTTTAC ACCGCCGACG GCAATATCCG TATGGGGACG TGGTATATGG
1601 CGGACACCAA ACGCCGCTG CAAAACAACG AAGTCCTCGC CACCGCAGGC
1651 TATAACGCCG GTCCCGGCAG GGCGCGCCGA TGGCAGGCGG ACACGCCCTT
1701 CGAAGGCGCG GTATATGCCG AAACCATCCC GTTTTCCGAA ACGCGCGACT
1751 ATGTCAAAAA AGTGATGGCC AATGCCGCT ACTACGCTC CCTCTTCGGC
1801 GCGCCGACA TCCCGCTCAA ACAGCGTATG GGCATTGTCC CCGCCGCTG
1851 A

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 019.a>:

```

a019.pep
1  MYPPSLKHSL PLLVXLVLA CSXTNTLSAD KTPADNIETA DLSASVPTXP
51  AEPEXKTXAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARRQWT LXAXEYAKLE PAXRAQEVEC YADSSRNDYT RAAELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWGV
251 GHYQSONLNV PAALDXXGKV ADRRLTDDQ IEWYARAAXX XRXXXXXXAXX
301 XXXXXXXXXX XXXXXXXXXAR SRAATGNTQX AXKLYKQAAA XGNFYAVLX
351 GEELGRXIDT RNNVPDAGKX SVLRMAEDGA IKRALVLFN SRTAGDAKMR
401 RXAQAERFRA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNYYL

```



451 RYISXXXDTV IRHAQNVNVD PAWVYGLIRQ ESRFVMGAQS RVGAQGLMQV  
501 MPATAREIAG KIGMDAAQLY TADGNIRMGY WYMADTKRRL QNNEVLATAG  
551 YNAGPGRARR WQADTPLEGA VYAETIPFSE TRDYVKKVMA NAAYYASLFG  
601 APHIPLKQRM GIVPAR\*

m019/a019 88.9% identity over a 524 aa overlap

m019.pep	10	20	30	40	50	60
	MYLPSMKHSLPLLAALVLAACSSNTNLPAGKTPADNIETADLSASVPTPAEPEPKTLAD					
a019	MYPPSLKHSLPLLXLVLAACSTNTLSADKTPADNIETADLSASVPTXPAEPEKXTXAD					
	10	20	30	40	50	60
m019.pep	70	80	90	100	110	120
	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKSLGARQWTLFAQEYAKLE					
a019	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKSLGARQWTLXAXEYAKLE					
	70	80	90	100	110	120
m019.pep	130	140	150	160	170	180
	PAGRAQEVECYADSSRNDYTRAAELVKNTGKLPSCGCTKLLEQAAASGLLDGNDAWRRVRG					
a019	PAXRAQEVECYADSSRNDYTRAAELVKNTGKLPSCGCTKLLEQAAASGLLDGNDAWRRVRG					
	130	140	150	160	170	180
m019.pep	190	200	210	220	230	240
	LLAGRQTTDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEMESGLSL					
a019	LLAGRQTTDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEMESGLSL					
	190	200	210	220	230	240
m019.pep	250	260	270	280	290	300
	EQRSFAWGVLGHYQSQNLNVPAALDYYGKVADRRQLTDDQIEWYARAALRARRWDELASV					
a019	EQRSFAWGVLGHYQSQNLNVPAALDYGKVADRRQLTDDQIEWYARAAXXXRXXXXXAXX					
	250	260	270	280	290	300
m019.pep	310	320	330	340	350	360
	ISHMPEKLQKSPTWLYWLARSRAATGNTQEAELKLYQAAATGRNFYAVLAGEELGRKIDT					
a019	XXXXXXXXXXXXXXXXXARSRAATGNTQXAKLYQAAAXGXNFYAVLXGEELGRXIDT					
	310	320	330	340	350	360
m019.pep	370	380	390	400	410	420
	RNNVPDAGKNSVRRMAEDGAVKRALVLFNSQSAGDAKMRRQAQAEWRFATRGFDEDKLL					
a019	RNNVPDAGKXSVLRMAEDGAIKRALVLFNSRTAGDAKMRRXAQAEWRFATRGFDEDKLL					
	370	380	390	400	410	420
m019.pep	430	440	450	460	470	480
	TAAQTAFDHGFYDMAVNSAERTDRKLNLYTLRYISPFKDTVIRHAQNVNVDPAWVYGLIRQ					
a019	TAAQTAFDHGFYDMAVNSAERTDRKLNLYTLRYISXXXDTVIRHAQNVNVDPAWVYGLIRQ					
	430	440	450	460	470	480
m019.pep	490	500	510	520		
	ESRFVIGAQSRVGAQGLMQVMPATAREIAGKIGMDAAQLYTADG					
a019	ESRFVIGAQSRVGAQGLMQVMPATAREIAGKIGMDAAQLYTADGNIRMGYWMADTKRRL					
	490	500	510	520	530	540
a019	QNNEVLATAGYNAGPGRARRWQADTPLEGAVYAETIPFSETRDYVKKVMAAAYYASLFG					
	550	560	570	580	590	600

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 019 shows 95.5% identity over a 89 aa overlap with a predicted ORF (ORF 019.ng) from *N. gonorrhoeae*:

g019/m019

		10	20	30	40	49
g019.pep		LLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPEGKTLAD				
m019		MYLPSMKHSLPLLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPEKTLAD				
		10	20	30	40	50
	50	60	70	80	89	
g019.pep	YGGYPSALDAVKQNNDAAAAYLENAGDSAMAENVRKEWL					
m019	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKSLGARQWTLFAQEYAKLE					
	70	80	90	100	110	120

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 111>:

g023.seq

```

1  ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GCGATTGGGT
51  AATGCAGCGT GCGACTGCGG TTATTATGTT GATTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTGCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATTT
151 TTTAGTCAAG CTTGGGTAAA AGTATTTACC CAAGTGAGCT TTATCGCCGT
201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCTTCGG CGTGCCTTTG TTTTTCAGG TTGCCACCAT TGctGGCTG
301 GTCGGCTGCC TCGTGATTTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 023.ng>:

g023.pep

```

1  MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFLAL PKEYPAWQAF
51  FSQAWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

m023.seq

```

1  ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GCGATTGGGT
51  GATGCAACGT GCGACTGCGG TTATTATGTT GATTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTTCCCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT
201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCTTCGG CGTGCCTTTG TTTTTCAGG TTGCCACCAT CGTTTGGCTG
301 GTCGGCTGTC TCGTGATTTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 023>:

m023.pep

```

1  MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFLSL PKEYSAWQAF
51  FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

a023.seq

```

1  ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GGGATTGGGC
51  GATGCAACGT GCGACCGCGG TTATTATGTT GATTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTGCTCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT

```

```

201  ATTCTTGAC  GCTTGGGTGG  GTATCCGCGA  TTTGTGGATG  GACTATATNA
251  AACCTTTCGG  CGTGCGTTTG  TTTTTCAGG  TTGCCACCAT  CGTCTGGCTG
301  GTCGGCTGCT  TGGTGTATTC  AATTAAAGTA  ATTTGGGGGT  AA

```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 023.a>:

```

a023.pep
1  MVERKLTGAH YGLRDWAMQR ATAVIMLIYT VALLVVLFA L PKEYSAWQAF
51  FSQTWVKVET QVSFIAVFLH AWVGIRDLWM DYXKPFVRL FLQVATIVWL
101 VGCLVYSIKV IWG*

```

m023/a023 96.5% identity over a 113 aa overlap

```

          10      20      30      40      50      60
m023.pep  MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFLSLPKEYSAWQAFFSQTWVKVFT
          |||||:|||||:|||||:|||||:|||||:|||||
a023      MVERKLTGAHYGLRDWAMQRATAVIMLIYTVALLVVLFLALPKEYSAWQAFFSQTWVKVFT
          10      20      30      40      50      60

          70      80      90     100     110
m023.pep  QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
          |||||:|||||:|||||:|||||:|||||
a023      QVSFIAVFLHAWVGIRDLWMDYXKPFVRLFLQVATIVWLVGCLVYSIKVIWGX
          70      80      90     100     110

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 023 shows 97.3% identity over a 113 aa overlap with a predicted ORF (ORF 023.ng) from *N. gonorrhoeae*:

g023/m023

```

          10      20      30      40      50      60
g023.pep  MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFLALPKEYPAWQAFFSQAWVKVFT
          |||||:|||||:|||||:|||||:|||||
m023      MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFLSLPKEYSAWQAFFSQTWVKVFT
          10      20      30      40      50      60

          70      80      90     100     110
g023.pep  QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
          |||||:|||||:|||||:|||||
m023      QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
          70      80      90     100     110

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 117>:

```

g025.seq
1  ATGTTGAAAC AAACgACACT TTTGGCAGCT TGTACCGCCG TTGCCGCTCT
51  GTTGGGCGGT TGcgCCACCC AACAGCCTGC TccTGTCATT GCAGGCAATT
101 CAGGTATGCA GACCGTATCG TCTGCGCCGG TTTACAATCC TTATGGCGCA
151 ACGCCGTACA ATGCCGCTCC TGCCGCCAac gatgcGCCgT ATGTGCCGCC
201 CGTGCAAact ggcgcggttT ATTCGCCTCC TGCTTATGTT CCGCcgTCTG
251 CACCTGCCGT TTCGGgtaca tatgtTCCTT CTTACGCACC CgtcgACATC
301 aacgCGGCGa cgCataCTAT TGTGCGTGGC GACACgGtgt acaACATTTc
351 caaAcgCtac CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
401 CCGACAATAC GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCaggA
451 TATGCCGCAC CGAAAACCGC AGCCGTAGAA AGCAGGCCCG CCGTACCGGC
501 TGCCCGCAA ACCCCTGTGA AAGCCGCCGC gcaACCGCCC GTTCAGTCCG
551 CGCCGCAACC TGCCCGCGCC GCTGCGGAAA ATAAAGCGGT TCCCGCCCCC
601 GCGCCCGCCC CGCAATCTCC TGCCGCTTCG CTTCCGGCA CGCGTTCGGT
651 CGGCGGCATT GTTTGGCAGC GTCCGACCCA AGGTAAAGTG GTTGCCGATT

```

```

701 TCGGCGGCGG CAACAAGGGT GTCGATATTG CCGGCAATGC CGGACAACCC
751 GTTTTGGCGG CGGCTGACGG CAAAGTGGTT TATGCCGGTT CAGGTTTGAG
801 GGGATACGGA AACTTGGTCA TCATCCAGCA CAATTCCTCT TTCCTGACCG
851 CGTACGGGCA CAACCAAAAA TTGCTGGTCG GCGAAGGTCA GCAGGTCAAA
901 CGCGGTGACG AGGTTGCTTT GATGGGTAAT ACCGATGCTT CCAGAACGCA
951 GCTTCATTTC GAGGTGCGTC AAAACGGCAA ACCGGTTAAC CCGAACAGCT
1001 ATATCGCGTT CTGA

```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 025.ng>:

```

g025.pep
1  MLKQTTLLAA CTAVAALLGG CATQQPAPVI AGNSGMQTVS SAPVYNPYGA
51  TPYNAAAPAN DAPYVPPVQT APVYSPPAYV PPSAPAVSGT YVPSYAPVDI
101 NAATHTIVRG DTVYNISKRY HISQDDFFRAW NGMTDNTLSI GQIVKVKPAG
151 YAAPKTAAVE SRPAVPAAAQ TPVKPAAQPP VQSAPQPAAP AAENKAVPAP
201 APAPQSPAAS PSGTRSVGGI VWQRPTQGV VADFGGNGK VDIAGNAGQP
251 VLAADGKVV YAGSGLRGY NLVIIQHNS FLTAYGHNQK LLVGEQQQVK
301 RGQQVALMGN TDASRTQLHF EVRQNGKPVN PNSYIAF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

```

m025.seq (partial)
1  ..GTGCCGCCG TGCAAAGCGC GCCGGTTTAT ACGCCTCCTG CTTATGTTCC
51  GCCGTCTGCA CCTGCCGTTT CGGGTACATA CGTTCCTTCT TACGCACCCG
101 TCGACATCAA CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC
151 AACATTTCCA AACGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGAA
201 CGGCATGACC GACAATACGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC
251 CGGCAGGATA TGCCGCACCG AAAGCCGCAG CCGTAAAAAG CAGGCCCGCC
301 GTACCGGCTG CCGCGCAACC GCCCGTACAG TCCGCACCCG TCGACATTAA
351 CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC AACATTTCCA
401 AACGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGAA CGGCATGACC
451 GACAATATGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC CGGCAGGATA
501 TGCCGCACCG AAAACCGCAG CCGTAGAAAG CAGGCCCGCC GTACCGGCTG
551 CCGTGCAAAC CCCTGTGAAA CCGCGCCGCG AACC GCCTGT GCAGTCCGCG
601 CCGCAACCTG CCGCGCCCGC TGCGGAAAAT AAAGCGGTTT CCGCGCCCGC
651 CCGCAATCT CCTGCCGCTT CGCCTTCCGG CACGCGTTCT GTGCGCGGCA
701 TTGTTTGGCA GCGTCCGACG CAAGGTAAAG TGGTTGCCGA TTTCGCGCGC
751 AACAAACAGG GTGTCGATAT TGCCGGTAAT GCGGGACAGC CCGTTTTGGC
801 GGCGGCTGAC GGCAAAGTGG TTTATGCCGG TTCAGGTTTG AGGGGATACG
851 GAACTTGGT CATCATCCAG CATAATTCTT CTTTCTGAC CGCATACGGG
901 CACAACCAA AATTGCTGGT CGGCGAGGGG CAGCAGGTCA AACGCGGTCA
951 GCAGGTTGCT TTGATGGGCA ATACCGATGC TTCCAGAACG CAGCTTCATT
1001 TCGAGGTGCG TCAAAACGGC AAACCGGTTA ACCCGAACAG CTATATCGCG
1051 TTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 025>:

```

m025.pep (partial)
1  ..VPPVQSAPVY TPPAYVPPSA PAVSGTYVPS YAPVDINAAT HTIVRGDTVY
51  NISKRYHISQ DDFRAWNGMT DNTLSIGQIV KVKPAGYAAP KAAAVKSRPA
101 VPAAQPPVQ SAPVDINAAT HTIVRGDTVY NISKRYHISQ DDFRAWNGMT
151 DNMLSIGQIV KVKPAGYAAP KTAAVESRPA VPAAVQTPVK PAAQPPVQSA
201 PQPAAPAAEN KAVPAPAPQS PAASPSGTRS VGGIVWQRPT QGKVVADFGG
251 NNKGVDIAGN AGQPVLAAD GKVVYAGSGL RGYGNLVIIQ HNSSFLLTAYG
301 HNQKLLVGEQ QQVKRGQOVA LMGNTDASRT QLHFEVRQNG KPVNPNSYIA
351 F*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```

a025.seq
1  ATGTTGACAC CAACAACACT TTAGGTAGCT TGTACCGCCC TTGCCGCTCA
51  GTTGGGCGGA TGCCCCACCC AACACCCTTC TCCTGTCATT GCAGGCAATT
101 CAGGTATGCA GACCGTACCG TCTGCGCCGG TTTACAATCC TTATGGCGCA

```

```
151 ACGCCGTACA ATGCCGCTCC TGCCGCCAAC GATGCGCCGT ATGTGCCGCC
201 GGTGCAAAGC GCGCCGGTTT ATANGCCTCC TGCTTATGTT CCGCCGTCTG
251 CACCTGCCGT TTCGGGTACA TACGTTTCCTT CTTACGCANC CGTCGACATC
301 AACGCGGCGA CCCATACTAT TGTGCGCGGC GACACCGTGT ACAAGATTTC
351 CAAATGCTAC CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
401 CCGACAATAC GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCAGGA
451 TATGCCGCAC CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC
501 TGCCGCGCAA CCGCTCGTAC AGTCCGCACC CGTCGACATC AACCGGGCGA
551 CGCATACTAT TGTGCGCGGC GACACGGTGT ACAACATTTC CAAACGCTAC
601 CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA CCGACAATAC
651 GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCAGGA TATGCCGCAC
701 CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC TGCCGTGCAA
751 ACCCGTGTGA AACCCGCCGC GCAACCGCCT GTGCAGTCCG CGCCGCAACC
801 TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCGCCC GCCCGCAAT
851 CTCCTGCGCG TTCGCCTTCC GGCACGCGTT CCGTCGGCGG CATGTGTTGG
901 CAGCGTCCGA CGCAAGGTAA AGTGGTTGCC GATTTCGGCG GCAACAACAA
951 GGGTGTGAT ATTGCAGGAA ATGCGGGACA GCGCGTTTTC GCGGCGGCTG
1001 ACGGCAAAGT GGTTTATGCA GGTTCCGGTT TGAGGGGATA CGGCAATTTG
1051 GTCATCATCC AGCATAATTC TTCCTTCCTG ACCGCATACG GGCACAACCA
1101 AAAATTGCTG GTCGGCGAAG GCCAGCAGGT CAAACGCGGG CAGCAGGTCG
1151 CTTTGATGGG CAATACCGAG GCTTCTAGAA CGCAGCTTCA TTTCGAGGTG
1201 CGGCAAAACG GCAAACCGGT TAATCCGAAC AGCTATATCG CGTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 025.a>:

```
a025.pep
  1 MLTP TTL*VA CTALAAQLGG CPTQHPSPIV AGNSGMQTVP SAPVYNPYGA
 51 TPNNAAPAA DAPYVPPVQS APVYXPPAYV PPSAPAVSGT YVPSYAXVDI
101 NAATHIVRG DTVYKISKCY HISQDDFRAW NGMTDNTLSI GQIVKVKPAG
151 YAAPKAAAVK SRPAVPAAQ PLVQSAPVDI NAATHIVRG DTVYNISKRY
201 HISQDDFRAW NGMTDNTLSI GQIVKVKPAG YAAPKAAAVK SRPAVPAAVQ
251 TPVKPAAQPP VQSAPQPAAP AAENKAVPAP APQSPAASPS GTRSVGGIVW
301 QRPTQGVVVA DFGGNNKGVD IAGNAGQPV LAAADGKVYA GSGLRGYGNL
351 VIIQHNSFL TAYGHNQKLL VEGEQQVKRG QQVALMGNTE ASRTQLHFEV
401 RQNGKPVNPN SYIAF*
```

m025/a025 97.4% identity over a 351 aa overlap

```
                                10      20      30
m025.pep                      VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
                                |||||:|||||
a025                          GMQTVPSAPVYNPYGATPYNAAPAANDAPYVPPVQSAPVYXPPAYVPPSAPAVSGTYVPS
                                40      50      60      70      80      90

                                40      50      60      70      80      90
m025.pep                      YAPVDINAATHIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
                                || |||||:||||| |||||
a025                          YAXVDINAATHIVRGDTVYKISKCYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
                                100     110     120     130     140     150

                                100     110     120     130     140     150
m025.pep                      KAAAVKSRPAVPAAQPPVQSAPVDINAATHIVRGDTVYNISKRYHISQDDFRAWNGMT
                                |||||
a025                          KAAAVKSRPAVPAAQPLVQSAPVDINAATHIVRGDTVYNISKRYHISQDDFRAWNGMT
                                160     170     180     190     200     210

                                160     170     180     190     200     210
m025.pep                      DNMLSIGQIVKVKPAGYAAPKTAAVESRPVPAVQTPVKPAAQPPVQSAPQPAAPAAEN
                                || |||||:||||| |||||
a025                          DNTLSIGQIVKVKPAGYAAPKAAAVKSRPAVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN
                                220     230     240     250     260     270

                                220     230     240     250     260     270
m025.pep                      KAVPAPQSPAASPSGTRSVGGIVWQRPTQGVVADFGGNNKGVDIAGNAGQPVLAAD
                                |||||
```

208

```

a025      KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAD
           280      290      300      310      320      330

m025.pep  GKVVYAGSGLRGYGNLVIIQHNSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDASRT
           280      290      300      310      320      330
|||||
a025      GKVVYAGSGLRGYGNLVIIQHNSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTASRT
           340      350      360      370      380      390

m025.pep  QLFHEVRQNGKPVNPNSYIAFX
           340      350
|||||
a025      QLFHEVRQNGKPVNPNSYIAFX
           400      410

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 025 shows 75.6% identity over a 353 aa overlap with a predicted ORF (ORF 025.ng) from *N. gonorrhoeae*:

m025/g025

```

m025.pep  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
           10      20      30
|||||
g025      GMQTVSSAPVYNPYGATPYNAAPAANDAPYVPPVQTAPVYSPPAYVPPSAPAVSGTYVPS
           40      50      60      70      80      90

m025.pep  YAPVDINAATHITVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
           40      50      60      70      80      90
|||||
g025      YAPVDINAATHITVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP
           100     110     120     130     140     150

m025.pep  KAAAVKSRPAVPAAAQPPVQSAPVDINAATHITVRGDTVYNISKRYHISQDDFRAWNGMT
           100     110     120     130     140     150
|
g025      K-----

m025.pep  DNMLSIGQIVKVKPAGYAAPKTAAVESRPAPVAAVQTPVKPAAQPPVQSAPQPAAPAAEN
           160     170     180     190     200     210
|||||
g025      -----TAAVESRPAPVAAQTPVKPAAQPPVQSAPQPAAPAAEN
           160     170     180     190

m025.pep  KAVPAPAP--QSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLA
           220     230     240     250     260
|||||
g025      KAVPAPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLA
           200     210     220     230     240     250

m025.pep  ADGKVVYAGSGLRGYGNLVIIQHNSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
           270     280     290     300     310     320
|||||
g025      ADGKVVYAGSGLRGYGNLVIIQHNSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
           260     270     280     290     300     310

m025.pep  RTQLHFEVRQNGKPVNPNSYIAFX
           330     340     350

```

g025                    |||||  
                       RTQLHFEVRQNGKPVNPNSYIAFX  
                       320                    330

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 113>:

g031.seq  
 1 ATGGTGTCCTCC TCCGCTTCAG ATTCGGCAAC CACTTTAAAC GCCGACATTC  
 51 TGACAATTTTC CTTTTCCGCC AGCCAAATAT CATGCGTATC TTTCGGTTCCG  
 101 GGCTTGTTGG GCATGGCAAC CTTCAACAGC CGCGCCATCA CAGGAATCGT  
 151 CGTTCCTTGA ATCAGCAGCG ACAGCACCAC CACGGCAAAC GCCACATCAA  
 201 ACAGCAGGTG CGAATTGGGA ACGCCCATCA CCAGCGGCAT CATCGCCAGC  
 251 GAAATCGGTA CGGCTCCTCG CAAGCCCAAC CAACTGATAT ACGCCTTTTC  
 301 ACGCAGGCTG TAATTGAATT TCCACAAACC GCCGAACACT GCCAGCGGAC  
 351 GCGCGACCAG CATCAGGAAC GCCGCAATCG CCAAGGCTTC CGCCGCCCTG  
 401 TCCAACACGC CGGCGGGAGA AACCAGCAGA CCGAGCATGA CGAACAAAGT  
 451 TGCTGCGCC AGCCAAGCCA AACCCTCCAT CACACGCAA ACCTGTTCCG  
 501 TcgACGGTT GCGCTGGTTA CCGACAATGA TGCCGGCAAG GTAAACCGCC  
 551 AAAAAAGCCG TGCCGCTAT GGTATTGGTA AACGCAAACA CAAGCAGCCC  
 601 GCCCGACACA ATCATCAGCG CGTACAGACC TTCCGtacac acctccaatt  
 651 cccaatcaac gtcatagtct tctcccggtg taaaatgttc ttcacttcag  
 701 aatccccccc ttcttcccag cccgaaacct tcatgtgtta naccctgggg  
 751 tgccccaacg gatttagtaa cctcccaatg actctgcttg tcgccccctt  
 801 cgcccgcttt ctccctccgg gaaaacttgt tgctcccgct ttacattaa

This corresponds to the amino acid sequence <SEQ ID 114; ORF 031.ng>:

g031.pep  
 1 MVSLRFRFGN HFKRRHSDNF LFRQPNIMRI FRFGLVGHGN LQQPRHHRNR  
 51 RSLNQQRQHH HGKRHIKQOV RIGNAHHQRH HRQRNRYGSS QAQPTDIRLF  
 101 TQAVIEFPQT AEHCQTRDQ HQERRNRQGF RRPVQHAGGR NQOTEHDEQS  
 151 CLRQPSQTVH HTQNVFRRTV ALVTDNDAGK VNRQKAAAAY GIGKRKHQKQ  
 201 ARHNHQRVQT FRTHLQFPIN VIAVSRVKMF FTSESPSSQ PETFMCXTLG  
 251 CPNGFSNLPM TLLVAPFARF LLPGKLVVPV LH\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

m031.seq (partial)  
 1 ...CGCCTGAAGC ACGGTGTCGG ACTGCATTTT TATTCGGCTA TACGCCTTTT  
 51 CACGCAGGCT GTAATTGAAT TTCCACAAAC CGCCGAACAC TGCCGACGGA  
 101 CGCGCGACCA GCATCAGGAA CGCCGCAATC GCCAAGGCTT CCGCCGCCCT  
 151 GTCCAACACG TTGGCAGGAG AAACCAGCAG CAAAGGCATT CCCAAACGTG  
 201 CGGACAAAGT GGTGAAACC ACGCTCAGAA ACAACAGTGC GCCACCCGGC  
 251 AG....

This corresponds to the amino acid sequence <SEQ ID 116; ORF 031>:

m031.pep (partial)  
 1 ...RLKHGVGLHF YSAIRLFTQA VIEFPQTAEH CRRTRDQHQE RRRNRQGFRRP  
 51 VQHVGRNRQQ QRHSQTCGQS GRNHAQKQQC ATRQ....

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

a031.seq  
 1 ATACGCCTTT TCACGCAGGC TGTAATTGAA TTTCCACAAA CCGCCGAACA  
 51 CTGCCGGCGG ACGCGCGACC AGCATCAGGA ACGCCGCAAT CGCCAAGGCT  
 101 TCCGCCGCCC CGTCCAACAC GTTGGCAGGA GAAACCAGCA GCAAAGGCAT  
 151 TCCCAAACGT GCGGACAAAG TGGTCGAAAC CACGCTCAGA AACAAACAGT  
 201 CGCCACCCGG CAG

This corresponds to the amino acid sequence <SEQ ID 118; ORF 031.a>:

a031.pep (partial)  
 1 IRLFTQAVIE FPQTAEHCRR TRDQHQERRN RQGFRFPVQH VGRNRQQQRH  
 51 SQTCGQSGRN HAQKQQCATR Q

Computer analysis of this amino acid sequence gave the following results:  
Homology with a predicted ORF from *N. gonorrhoeae*  
 ORF 031 shows 60.0% identity over a 85 aa overlap with a predicted ORF (ORF 031.ng)  
 from *N. gonorrhoeae*:  
 m031/g031

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 119>:

This corresponds to the amino acid sequence <SEQ ID 110; ORF 032.ng>:

BNSDOCID: <WO\_\_9957280A2\_1\_>



```

101 EQRVVAHRQR VAAVHGQIQH PVQPFRLQGF GYALGLLRRF DVGGRVGAHQ
151 PAFDQPGAIL PPRRLARQR PTVQTALRQP PQRRLKIAPR QVLRHAACIF
201 RRHLCQCKQ FFQIAPVCRN RVLRLALAH VFIQSVKIRR KPVQNHNRPT
251 QISKNQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```

m032.seq (partial)
  1 ATGCGGCGAA ACGTGCmTGC mGTGCGCGTT kTGCGCCGCC CATTGCGCCA
 51 AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGACCAGCGG CACCTGCCGC TgTT.GCGCC
201 CTTTGCCGAT AAcGTGTACC CACGcyTCGT CCAAATAGAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAA CCGTGCCGT ACATCAGTTC
301 GAACAGGGCG TGGTCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTTCAGC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG CGTGCAATCAG
451 GCCGCGCTTT ACCAGCCAA CGCAATACTG CCGCCAAGAC GAAAGCTTGC
501 GAGCCAGCGT CCGTTCCCCC AAACCGCG...

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 032>:

```

m032.pep (partial)
  1 MRRNVXAVAV XRRPLRQTFD DLALAQARAV PAGKQGFAGR CRLTQRQIVF
 51 QGFHAFADQR HLPLXAPFAD NVYPRXVQID IICIQAVYLA HAQTAHVHOF
101 EQGVVAHRQR VAAVHGQIQH PVQPFRLQGF GYALGLLRRF DVGGRVGVHQ
151 AALYQPNAIL PPRRLASQR PFPQTA...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

```

a032.seq
  1 ATGCGGCGAA ACGTGCCTGC CGTCGCGGTT TTGCGCCGCC CATTGCGCCA
 51 AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGCCGC TGCTTGCGTC
201 CTTTGCCGAT AAcGTGTACC CACGCTCGT CCAAATATAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAA CCGTGCCGT ACATCAGTTC
301 GAACAGCGCG TGATCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTTCAGC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG TATGCAGCAG
451 ACCGCGTTG ACCAGCCAGG CGCAATACTG CCGCCAAGAC GACAGCTTGC
501 GCGCCAGCGT CCGCGCATTC AAACCGCGCT GCGACAGCCG CCGCAACGCC
551 GCCGTAAAAT CGCGCTGCGA CAAGCCCTGC GGCACGCCGC CTGCATCTTC
601 AGACGGCATT TGTGCCAACA GCGCAAACAG TTCTTCCAAA TCGCGCCGGT
651 ATGCCGCCAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAAA
701 TAAGCGTCAA AATGCGCCGC AAACCCGTCC AAAACCATAA CCGCCCCACA
751 CAAATATCAA AAAACAGTG A

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 032.a>:

```

a032.pep
  1 MRRNVPAAVAV LRRPLRQTFD DLALAQARAV PAGKQGFAGR CRLTQRQIVF
 51 QGFHAFAGQR NLPLASFAG NVYPRVQIY IICIQAVYLA HAQTAHVHOF
101 EQRVIAHRQR VAAVHGQIQH PVQPFRLQGF GYALGLLRRF DVGGRVGMQQ
151 TAFDQPGAIL PPRRLARQR PRIQTALRQP PQRRLKIALR QALRHAACIF
201 RRHLCQQRKQ FFQIAPVCRH RVLRLALAH VFIQSVKMRR KPVQNHNRPT
251 QISKKQ*

```

m032/a032 88.1% identity over a 176 aa overlap

```

          10      20      30      40      50      60
m032.pep  MRRNVXAVAVXRRPLRQTFDLALAQARAVPAGKQGFAGVRCRLTQRQIVFQGFHAFADQR
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a032      MRRNVPAAVAVLRRPLRQTFDLALAQARAVPAGKQGFAGVRCRLTQRQIVFQGFHAFAGQR
          10      20      30      40      50      60

```

Homology with a predicted ORF from *N. gonorrhoeae*

m032/g032

		10	20	30	40	50	60
m032.pep		MRRNVXAVAVXRRPLRQTFLDLALAQARAVPAGKQGFVAVRCRLTQRQIVFQGFHAFADQR					
		:					
g032		MRRNVPAVAVLRRPRFEAFDLALAQARAVPAGKQGFVAVRCRLTQRQIVFQGFHAFAGQR					
		10	20	30	40	50	60
		70	80	90	100	110	120
m032.pep		HLPLXAPFADNVYPRXVQIDIICIQAVYLAAHQTAAVHQQFEQGVVAHRQRVAAVHGGIQH					
		:					
g032		NLTLLAPFAGNVYPRFVQIYIICIQAVYLAAHQTAAVHQQLEQRVVAHRQRVAAVHGGIQH					
		70	80	90	100	110	120
		130	140	150	160	170	
m032.pep		PVQPFRLRQGFYALGLLRRFDVGGRGVGHQAALYQPNAILPPRRKLASQRPPFPOTA					
		:					
g032		PVQPFRLRQGFYALGLLRRFDVGGRGVGAHQPAFDQPGAILPPRRQLARQRPTVQTALRQP					
		130	140	150	160	170	180
g032		PQRRRKIAPRQVLRHAACIFRRHLCCQCKQFFQIAPVCRNRVLRLLALAHDFQISVKIRR					
		190	200	210	220	230	240

g033.seq

1	ATGGCGGCGG	CGGACAAACT	CTTGGGCGGC	GACCGCCGCA	GCGTCGCCAT
51	CATCGGAGAC	GGCGCGATGA	CGGCGGGGCA	GGCGTTTGAA	GCCTTGAATT
101	GCGCGGGCGA	TATGGATGTG	GATTTGCTGG	TCGTCTCAA	CGACAACGAA
151	ATGTCGATT	CCCCAACGT	CGGCGCGTTG	CCCAAATATC	TTGCCAGCAA
201	CGTCGTGCGC	GATATGCACG	GACTGTTGAG	TACCGTCAA	GCGCAAAcgg
251	GCAAGGTATT	AGACAAAAT	CCCGGCGCGA	TGGagtTTGC	CCAAAAAGTC
301	GAACAcaaa	TCAAAACCCT	TGCCGAAGAA	GCGGAACACG	CCAAACAGTC
351	GCTGTGCTG	TTTGAAAATT	TCGGCTTCCG	CTACACCGGC	CCCTGGACG
401	GACACAACGT	CGAGAACTTG	GTGGACGTAT	TGAAAGACTT	GCGCAGCCGC
451	AAAGGCCCTC	AGTTGCTGCA	CGTCATCACC	AAAAAGGGCA	ACGGCTACAA
501	ACTCGCCGAA	AACGACCCcg	tcaAATACCA	CGCCGTCGcc	aACCTGCcta
551	AAGAAGGCGG	GGCGCAAAatg	cGTCGTGAAA	AAGAACCCAA	GCCCGCCgCc
601	aaaccgACCT	ATACCCAAGT	ATTCCGCCAAA	TGGCTGTGCG	ACCGGGCAGC
651	GGCAGATTCC	CGACTGGTTG	CGATTACCCC	CGCCATGCGC	GAGGGCAGCG

```

701 GACTGGTGGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT tacCTTTTGCC GGCGGTTTGG CGTGCGAAGG
801 CATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGCC
901 GTCGACCGTG CGGGCATCGT CGGCGCGGAC GGTCCGACCC ATGCCGGCTT
951 GTACGATTTG AGCTTCTTGC GCTGTGTGCC GAACATGATT GTTGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGCTGC TTTCGACCTG CTATCAGGCG
1051 GATGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101 GGTTTCAGAC GGCATGGAAA CCGTGGAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCC TTcatTGCCT TCGGCAGTAT GGTCGCCACC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTt
1251 cgtcaaacCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCAcg
1301 accGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTCTTGG AAGTGTGGC GAAACACGGC ATCTGCAAC CCGTTTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGAACA CGGCGATCCG AAAAACTTT
1451 TGGACGATTT GGGTTTGAGT GCCGAAGCGG TGGAACGCCG GGTGCGCGAG
1501 TGGCTGCCCG ACCGTGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 033.ng>:

**g033.pep**

```

1 MAAADKLLGG DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTAAE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDLKDLRSR
151 KGPQLLHVIT KKGNGYKLA E NDPVKYHAVA NLPKEGGAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAT
401 ALAVAELKNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKHG ICKPVLLLG V ADTVTEHGDP KKLLDDLGLS AEAVERVRRE
501 WLPDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

**m033.seq**

```

1 ATGGCGGCGG CAGACAAACT CTTGGGCAGC GACCGCCGCA GCGTCGCCAT
51 CATCGGCGAC GGCGCGATGA CGGCGGGGCA GGCGTTTGAA GCCTTGAATT
101 CGCcaG.CGA TATGGATGTr GATTTGCTrG TCGTCTCAA CGACAACGAA
151 ATGTCGATTT CCCCCAACGT CGGCGCGCTG CCGAAATACC TTGCCAGCAA
201 CGTCGTGCGC GATATGCACG GCCTGTTGAG TACCGTCAA GCGCAAACGG
251 GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CAAAAAGTC
301 GAACACAAAA TCAAAACCCT TGCCGAAGAA GCCGAACAG CCAAACAGTC
351 GCTGTCTTTG TTTGAAAAC TCGGCTTCCG CTACACCGGC CCCGTGGACG
401 GACACAACGT CGAAAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC
451 AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501 ACTCGCCGAA AACGACCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA
551 AAGAAAGCGC GGCGCAAATG CCGTCTGAAA AAGAACCCTA GCCGCGCGCC
601 AAACCGACCT ATACCCAAGT GTTCGCAAAA TGGCTGTGCG ACCGGGCGGC
651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT TACCTTTGCC GGCGGTTTGG CTTGCGAAGG
801 GATGAAGCCC GTCGTGCGCA TTTATTCCAC CTTTTTACAA CGCGCCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTACCCGT TTTGTTTGCC
901 GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTCT
951 GTACGATTTG AGCTTTTTCG GTCGCGTGCC GAACATGATT GTCGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGTTGC TTTCGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101 GGTTTCAGAC GGCATGGAAA CCGTGGAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCGCC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG

```

```

1301 ACCGCATCGT TACCCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTGCTGG AAGTATTGGC GAAACACGGC ATCTGCAAAC CCGTTTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGAACGGCG TGTGCGCGCG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 033>:

m033.pep

```

1 MAAADKLLGS DRRSVAIIGD GAMTAGQAFE ALNCAXDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLDLRSR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVLF
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLSLTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
401 ALAVAELKNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKHG ICKPVLLLVG ADTVTGHGDP KKLLDDLGLS AEAVERVRRA
501 WLSDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

a033.seq

```

1 ATGGCGGCGG CGGACAAACA GTTGGGCAGC GACCGCCGCA GCGTCGCCAT
51 CATCGGCGAC GCGCGCATGA CGGCGGGTCA GCGCTTTGAA GCCTTGAAC
101 GCGCGGCGCA TATGGATGTG GATTGCTGG TCGTCCTCAA CGACAACGAA
151 ATGTCGATTT CCCCCAACGT CCGTGCCTG CCCAAATACC TTGCCAGCAA
201 CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAA GCGCAAACGG
251 GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTGC CAAAAAGTC
301 GAACATAAAA TCAAAACCCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC
351 ACTGTCTTTG TTTGAAAAC TCGGCTTCCG CTATACCGGC CCCGTGGACG
401 GACACAACGT CGAAAATCTG GTCGATGTAT TGGAAGACCT GCGCGGACGC
451 AAAGGCCCGC AGCTTCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501 ACTCGCCGAA AACGATCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA
551 AAGAAAGCGC GCGCGAAATG CCGTCTGAAA AAGAACCCTA GCCCGCCGCC
601 AAACCGACCT ATACCCAAGT GTTCGGCAA TGGCTGTGCG ACCGGGCGGC
651 GGCAGATTCC CGACTGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT TACCTTTGCC GCGGTTTGG CTTGCGAAGG
801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAA ACCTGCCCGT TTTGTTTGGC
901 GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTTT
951 GTACGATTTA AGCTTTTTGC GCTGCATTCC GAATATGATT GTCGCCGCGC
1001 CGAGCGATGA AAATGAATGC CGCCTGCTGC TTTCGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGTGCC
1101 GGTTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCCT
1201 GCATTGGCGG TCGCCGGAAG ACTGAACGCC ACCGTCGCCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG
1301 ACCGCATCGT TACCCCTGAA GAAAACGCCG AACAGGGCGG CGCAGGCAGC
1351 GCGGTGCTGG AAGTGTGTCG GAAACACGGC ATCTGCAAAC CCGTCTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGAACGGCG TGTGCGCGCG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 120; ORF 033.a>:

a033.pep

```

1 MAAADKQLGS DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLEDLRRR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVLF

```

301 VDRAGIVGAD GPTHAGLYDL SFLRCIPNMI VAAPSDENEC RLLSTCYQA  
351 DAPAAVRYPR GTGTGVPVSD GMETVEIGKG IIRREGKTA FIAFGSMVAP  
401 ALAVAGKLN TVADMRFKVP IDEELIVRLA RSHDRIVTLE ENAEQGGAGS  
451 AVLEVLAKHG ICKPVLLLG ADTVTGHGDP KKLLDDLGLS AEAVERRVRA  
501 WLSRDAAN\*

m033/a033 98.4% identity over a 509 aa overlap

m033.pep	10	20	30	40	50	60
a033	10	20	30	40	50	60
m033.pep	70	80	90	100	110	120
a033	70	80	90	100	110	120
m033.pep	130	140	150	160	170	180
a033	130	140	150	160	170	180
m033.pep	190	200	210	220	230	240
a033	190	200	210	220	230	240
m033.pep	250	260	270	280	290	300
a033	250	260	270	280	290	300
m033.pep	310	320	330	340	350	360
a033	310	320	330	340	350	360
m033.pep	370	380	390	400	410	420
a033	370	380	390	400	410	420
m033.pep	430	440	450	460	470	480
a033	430	440	450	460	470	480
m033.pep	490	500	510			
a033	490	500	510			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 033 shows 98.4% identity over a 509 aa overlap with a predicted ORF (ORF 033.ng) from *N. gonorrhoeae*:

m033/g033

m033 . pep	MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSISPNVGAL	60
g033	MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSISPNVGAL	60
m033 . pep	PKYLASNVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEAEHAHQSLSL	120
g033	PKYLASNVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEAEHAHQSLSL	120
m033 . pep	FENFGFRYTGPDVGHNVENLVDVLKDLRSRKGPPQLLHVITKKGNNGYKLAENDPVKYHAVA	180
g033	FENFGFRYTGPDVGHNVENLVDVLKDLRSRKGPPQLLHVITKKGNNGYKLAENDPVKYHAVA	180
m033 . pep	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRSLVAITPAMREGSGLVEFEQ	240
g033	NLPKEGGAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRSLVAITPAMREGSGLVEFEQ	240
m033 . pep	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA	300
g033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA	300
m033 . pep	VDRAGIVGADGPTHAGLYDLSFLRCVPMIVAAPSDENECRLLLSTCYQADAPAAVRYPR	360
g033	VDRAGIVGADGPTHAGLYDLSFLRCVPMIVAAPSDENECRLLLSTCYQADAPAAVRYPR	360
m033 . pep	GTGTGAPVSDGMETVEIGKGIIRREGKTAFFAGSMVAPALAVAELNATVADMRFVKP	420
g033	GTGTGAPVSDGMETVEIGKGIIRREGKTAFFAGSMVATALAVAELNATVADMRFVKP	420
m033 . pep	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTGHGDP	480
g033	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTEHGDP	480
m033 . pep	KKLLDDLGLSAEAVERRVRAWLSDRDAANX	510
g033	KKLLDDLGLSAEAVERRVREWLPDRDAANX	510

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 121>:

g034 . seq

1	ATGAGCCGTT	TATGGTTTTT	TGCCGTAAAA	AACATTATAA	TCCGCCTTAT
51	TTACCTATTG	CCCAAGGAGA	CACAAATGGC	ACTCGTATCC	ATGCGCCAAC
101	TGCTTGACCA	CGCCGCCGAA	AACAGCTACG	GCCTGCCCGC	GTTCAACGTC
151	AACAACCTCG	AACAAATGCG	CGCCATTATG	GAAGCCGCCG	ACCAAGTCAA
201	CGCGCCCGTC	ATCGTACAGG	CGAGCGCAGG	TGCGCGCAAA	TACGcgGCG
251	CGCCGTTTTT	GCGCCACCTG	ATTCTGGCGG	CAGTCGAAGA	ATTTCCGCAC
301	ATCCCGCTCG	TGATGCACCA	AGACCACGGC	GCATCGCCCG	ACGTgtGCCA
351	ACGCTCCATC	CAACTGGGCT	TCTCCTCCGT	GATGATGGAC	GGCTCTTTGC
401	TCGAAGACGG	CAAAACCCCT	TCTTCTTACG	AATACACGT	CAACGCCACC
451	CGTACCGTCG	TCAACTTCTC	CCACGCCTGC	GGCGTGTCGG	TCGAAGGCGA
501	AATCGGCGTA	TTGGGCAACC	TCGAAACCGG	CGAAGCAGGC	GAAGAAGACG
551	GAGTGGGCGC	GGCAGGCAAA	CTCTCACACG	ACCAAATGCT	CACCAGCGTT
601	GAAGATGCCG	TGCGTTTCGT	TAAAGATACC	GGCGTTGACG	CATTGGCGAT
651	TGCGCTCGGC	ACCAGCCACG	GCGCATACAA	ATTCAACCGT	CCGCCACAG

```

701 GCGACGTATT GCGTATCGAC CGCATCAAGG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGgctCCAGC TCCGTTCCGC AAGAatgGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG CAAAGTCAAC
901 ATCGATACCG ACCTGCGCCT CGCTTCCACC GCGCGGTAC GCCGTACCT
951 TGCCGAAAAC CCGTCCGACT TTGATCCGCG CAAATACTTG GGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCTCGACC GTTATCTTGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAT CAAACCTGTT TCGTTGGAAA AAATGGCAAG
1101 CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 122; ORF 034.ng>:

g034.pep

```

1 MSRLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCQSI QLGFSSVMMD GSLLLEDGKTP SSYEYNVNAT
151 RTVVNFSSHAC GVSVEGEIGV LGNLETGEAG BEDGVGAAGK LSHDQMLTSV
201 EDAVRFVKDT GVDALAI AVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKV NEYGGNIGET YGVPVEEIVE GIKHGVKRVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL GKTIEAMKQI CLDRYLAFGC
351 EGQAGKIKPV SLEKMASRYA KGELNQIVK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 123>:

m034.seq (partial)

```

1 ATGAGCTGTT TATGGTTTTT TGCTGTAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA WACAGCTACG GCyTGCCGCG GTTCAACGTC
151 AACAACCTCG WACAGATGCG CGCCATCATG GAGGCTGCAG ACCAAGTCGA
201 CGCCCCCGTC ATCGTACAGG CGAGTGCCGG TGCGCGCAAA TATGCGGGTG
251 CGCCGTTTTT ACGCCACCTG ATTTTGCGCG CTGTGGAAGT ATTTCCACAC
301 ATCCCCGTCG TCATGCACCA AGACCACGGC GCATCACC CGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TCTCCTCTGT AATGATGGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTACG AATACAACGT CAACGCCACA
451 CGTACCGTGG TTAACCTCTC CCACGCTTGC GCGGTATCCG TTGAAGGCCA
501 AATCGGCGTA TTGGGCAACC TCGAAACCGG CGATGCAGGC GAAGAAGACG
551 GTGTAGGCGC AGTGGGCAAA CTTTCCACG ACCAAATGCT GACCAGCGTC
601 GAAGATGCCG TATGTTTCGT TAAAGATACC GCGGTTGACG CATTGGCTAT
651 TGCCGTCGGC ACCAGCCACG GCGCATACAA ATTACCCCGT CCGCCCACAG
701 CGGATGTATT ACGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA C...

```

This corresponds to the amino acid sequence <SEQ ID 124; ORF 034>:

m034.pep (partial)

```

1 MSCLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE XSYGLPAFNV
51 NNLXQMRAIM EAADQVDAPV IVQASAGARK YAGAPFLRHL ILAAVEVFPH
101 IPVVMHQDHG ASPDVCQSI QLGFSSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFSSHAC GVSVEGEIGV LGNLETGDAG EEDGVGAVGK LSHDQMLTSV
201 EDAVCFVKDT GVDALAI AVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMH...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 125>:

a034.seq

```

1 ATGAGCCGTT TATGGTTTTT TGCCGCAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA AACAGCTACG GCCTGCCCGC GTTCAACGTC
151 AACAACCTCG ACAAATGCG CGCCATTATG GAAGCCGCCG ACCAAGTCAA
201 CGCGCCCGTC ATCGTACAGG CGAGCGCAGG TGCGCGCAAA TACGCGGGCG
251 CGCCGTTTTT GCGCCACCTG ATTTTGCGCG CTGTGGAAGA ATTTCCGAC
301 ATCCCCGTCG TGATGCACCA AGACCACGGC GCATCGCCCG ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TTTCTTCCGT GATGATGGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTATG AATACAACGT CAACGCCACC

```

```

451 CGTACCGTGG TTAATTTCTC CCACGCCTGC GCGGTATCCG TTGAAGGCGA
501 AATCGGCGTA TTGGGCAACC TCGAACTGG CGAAGCCGGC GAAGAAGACG
551 GTGTAGGCGC AGTGGGCAAA CTTTCCACG ACCAAATGCT CACCAGCGTC
601 GAAGATGCCG TGCGTTTCGT TAAAGATACC GCGGTGACG CATTGGCGAT
651 TGCCGTCGCG ACCAGCCACG GCGGTACAA ATTACCCGT CCGCCCACAG
701 GCGACGTGTT GCGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGGCTCCAGC TCCGTTCCGC AAGAATGGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG TAAAGTCAAC
901 ATCGATACCG ACTTGCGCCT TGCTTCCACC GCGCGGTAC GCCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TCGATCCGCG CAAATATTG AGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCTCGACC GCTACCTCGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAAT CAAACCGGTT TCCTTGAAA AAATGGCAAA
1101 CCGTTATGCC AAGGGCGAAT TGAACCAAT CGTCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 126; ORF 034.a>:

```

a034.pep
  1 MSRLWFFAAK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCQRSI QLGFSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFHAC GVSVEGEIGV LGNLETGEAG EEDGVGAVGK LSHDQMLTSV
201 EDAVRVFKDT GVDALAIAGV TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKV NEYGGNIGET YGVPVEEIVE GIKHGVKRVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL SKTIEAMKQI CLDRYLAFGC
351 EGQAGKIKPV SLEKMANRYA KGELNQIVK*

```

m034/a034 96.9% identity over a 257 aa overlap

```

m034.pep      10      20      30      40      50      60
MSCLWFFAVKNIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM
|| |||||:|||||
a034          10      20      30      40      50      60
MSRLWFFAAKNIIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNNLEQMRAIM

m034.pep      70      80      90      100     110     120
EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEVFPHIPVVMHQDHGASPDVCQRSI
|||||:|||||
a034          70      80      90      100     110     120
EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI

m034.pep     130     140     150     160     170     180
QLGFSSVMDGSLMEDGKTPSSYEYNVNATRTVVNFHACGVSVEGEIGVLGNLETGDAG
|||||:|||||
a034         130     140     150     160     170     180
QLGFSSVMDGSLMEDGKTPSSYEYNVNATRTVVNFHACGVSVEGEIGVLGNLETGEAG

m034.pep     190     200     210     220     230     240
EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAIAGVTSYGAYKFTRPPTGDVLRID
|||||:|||||
a034         190     200     210     220     230     240
EEDGVGAVGKLSHDQMLTSVEDAVRFVKDTGVDALAIAGVTSYGAYKFTRPPTGDVLRID

m034.pep     250
RIKEIHQALPNTHIVMH
|||||
a034         250     260     270     280     290     300
RIKEIHQALPNTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVKRVN

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 034 shows 96.5% identity over a 257 aa overlap with a predicted ORF (ORF 034.ng) from *N. gonorrhoeae*:



## m034/g034

m034.pep	MSCLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNXLQMRIM	60
g034	MSRLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNLEQMRIM	60
m034.pep	EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEVFPHIPVVMHQDHGASPDVCQRSI	120
g034	EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI	120
m034.pep	QLGFSSVMDGSLMEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGDAG	180
g034	QLGFSSVMDGSLLEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGEAG	180
m034.pep	EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAIAGVTSHGAYKFRPPTGDVLRID	240
g034	EEDGVGAAGKLSHDQMLTSVEDAVRFVKDTGVDALAIAGVTSHGAYKFRPPTGDVLRID	240
m034.pep	RIKEIHQALPNTHIVMH	257
g034	RIKEIHQALPNTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVRKVN	300

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 127>:

## g036.seq

```

1  ATGCTGAAGC CGTGTGTTGGT ATACAGTGCC TGTGCGGCGG cgttgCCTGC
51  GCGGACTTCG AGCAGCAGGC GTTGCCTGCC TTCGGGCAGA TGTGCGTACC
101 AATATTCGAG CAGGGCGGAC GCAACGCCCC GTCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCTGCCGT CTTTTTCCGC AAGGAAAACC TGTTCCGACG
251 GCGAAACAAG CGCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301 CAGACGGTAT CGAGCGCGGC CAGTGCAGCG CAGTCGGACG GTGAGGCTGG
351 GCGGATGTTT ATGTTCTGTC CTTCCGTTCC GCCTGTTCTT TGGCAGTCAG
401 GCGGATTTTG TTGCGGACGT AGAGCAGTTC GGCGTGTGCC GCGCCAGTTG
451 GCGGATAGCC GCCGCCGAGG GCGAGCGCGA GAAATCGGC GCGCGTCGGC
501 ATATCGGGTT TGCCTGAGAA GGGCGGACGG TTTTCCAGTG CGAACGCACT
551 GCCGATGCCG TCTGAAAAGA CGTACCCCTC GGGGAGGGCA ATGTCTGCCG
601 CCCTACCGAC TTGATAATCG CTCAAACGGC GGCGGTTTCA CGTGTGGAAC
651 CACGCATAAA AACTTTCGCC CATACGCGCG TCCGCAGCGG CGAGTATGCA
701 GCTTTGCGGC GCGGCGAGCG AGGCGGCGGC ATCGAGCGTG GGGATGCCGA
751 TTAAAGGCGT GTCGAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801 ATACGCAGTC CGGTAA

```

This corresponds to the amino acid sequence <SEQ ID 128; ORF 036.ng>:

## g036.pep

```

1  MLKPCLVYSA CAAALPARTS SSRRCVPSGR CAYQYSSRAD ATPRRRHSGA
51  VAIRCSSDSS GRFCQTIKAA ILPSFSARKT CSDGETSADS NWRCVHADGL
101 QTVSSAASAA QSDGEAGRMF MFVPSVPPVL WQSGRFCCGR RAVRRVPRQL
151 RDSRRRGRAR ENRRRSAYRV CLRRADGFPV RTHCRCLKR RTPRGGQCLP
201 PYRLDNRSNG GGSACRTHK TLRPYARPQR RVCSFAAAAA RRRHRAWGCR
251 LKACRTALPN LAPRRCRYAV R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 129>:

## m036.seq

```

1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTGCTGTC
51  ACGGACTTCG AGCAGCAGGC GTTGCCTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCCg.CGT CTTTTTCCGC AAGGAAAACC TGTTCCGACG
251 GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
301 CAGACGGCAT CGAGTGCAGC CAGCTCCTCA CAATCGGCAC AAACGGCAGC

```

```

351 GCGGATGTTT ACAGGGCGCGC TCTCCGTTTC GCCTGTTCTT TGGCAGTCAG
401 GGCGATTTTG TTGCGGACGT AGAGCAAACC GGCGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA AACTTTCGCC CATAAGAGCG TCCGTAGCGG CAAGGATGCA
701 GCTTTGCGGC GCGGCGAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAGGGGT ATCAAACGGC GTTGCCAAAC CCTGAGCTAC ACCGATGCCG
801 ATACGCAGTC CGGTAA

```

This corresponds to the amino acid sequence <SEQ ID 130; ORF 036>:

**m036.pep**

```

1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPXSFSAKT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPWLPMR ESRRQSAYPV CLRTAELLPA RTRCLRLKR RIPPAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*R QGCSFAAAAA RRRHRARVRR
251 LRGYQTALPN PELHRCRYAV R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 131>:

**a036.seq**

```

1 ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTGCTCTGC
51 ACGGACTTCG AGCAGCAGGC GTTGCCTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GCGGCGGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCT GGCAGTTCTT GCCAAACGAT
201 AAAGGCGGCA ATCCGCGCGT CTTTTTCCGC AAGGAAAACC TGTTCGGACG
251 GCGAAACCAG TCGGGAATCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301 CAGACGGCAT CGAGCGCGGC GAGTGCAGCG CAATCGGCAT AAACGGCGCG
351 GCGGATGTTT ACAGGCGCGC CCTCCGTTCC GCCTGTTCTT TGGCAGTCAA
401 GGCGATTTTG TTGCGGACGT AGAGCAGCTC GGCGTGTGCC GCAGCGACGG
451 CGGGAAAACC GCCTTCAGCC GCCAGATTGA GGAAGTCGGC GCGGTCGGC
501 ATATCGGGTT TGCCTGAGAA GGGCGGACGG TTTTCCAGCG CGAACGCATT
551 GCCGATGCCG TCTGAAAAGG CGCATCCTTC CGGCAGCCGG ATGTCTGCCG
601 CCCGACCGAC CTGATAATCG CTCAAACGGC GCGGTTTCAG CGTGTGCAAC
651 CATGCATAAA AACTTTCGCC CATAAGTGGC TCCGAGCGCG CAAGGATGCA
701 GCTTTGCGGC GCGGCGAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAGGAGT ATCAAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801 ATACGCAGTC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 132; ORF 036.a>:

**a036.pep**

```

1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPPSFSAKT CSDGETSADS NWRCVHADGL
101 QTASSAASAA QSA*TARRMF TGAPSVPPVL WQSRFFCCGR RAARRVPQRR
151 RENRLQPPD* GSRRRSAYRV CLRRADGFPA RTHCRCLKR RILPAAGCLP
201 PDRPDNRSNG GGSACRTMHK TLRPYVRPQR QGCSFAAAAA RRRHRARVRR
251 LKEYQTALPN LAPRRCRYAV P*

```

**m036/a036 85.6% identity over a 270 aa overlap**

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAIIPWRRHSGAVAIRCSSDSS					
a036	MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAIIPWRRHSGAVAIRCSSDSS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m036.pep	GRFCQTIKAAIPXSFSAKTCS DGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF					
a036	GRFCQTIKAAIPPSFSAKTCS DGETSADSNWRCVHADGLQTASSAASAAQSAXTARRMF					
	70	80	90	100	110	120

	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA					
a036	TGAPSVPPVLWQSGRFCCGRRARRVPQRRRENRLQPPDXGSRRRSAYRVCLRRADGFPA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERXPXRGCSFAAAAA					
a036	RTHCRCLKRRILPAAGCLPPDRPDNRSNGGGSACRTMHKTLRPYVRPQXRGCSFAAAAA					
	190	200	210	220	230	240
	250	260	270			
m036.pep	RRRHRARVRRRLRGYQTALPNPELHRCRYAVRX					
a036	RRRHRARVRRRLKEYQTALPNLAPRRCRYAVPX					
	250	260	270			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 036 shows 74.9% identity over a 271 aa overlap with a predicted ORF (ORF 036.ng) from *N. gonorrhoeae*:

m036/g036

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAIWRRHSGAVAIRCSSDSS					
g036	MLKPCLVYSACAAALPARTSSSRRCVPSGRCAYQYSSRADATPRRHSGAVAIRCSSDSS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m036.pep	GRFCQTIKAAIPXSFSARKTCSGGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF					
g036	GRFCQTIKAAILPSFSARKTCSGGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF					
	70	80	90	100	110	120
	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA					
g036	MEVPSVPPVLWQSGRFCCGRRARRVPRQLRDSRRRGRARENRRRSAYRVCLRRADGFVP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERXPXRGCSFAAAAA					
g036	RTHCRCLKRRTPRGGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPPRRVCSFAAAAA					
	190	200	210	220	230	240
	250	260	270			
m036.pep	RRRHRARVRRRLRGYQTALPNPELHRCRYAVRX					
g036	RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 133>:

m036-1.seq

```

1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTTGCCTGC
51  ACGGACTTCG AGCAGCAGGC GTTGCCTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GGCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCCGCCGT CTTTTCCTCG AAGGAAACC TGTTCCGACG
251 GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTG
301 CAGACGGCAT CGAGTGGCGC CAGCTCCTCA CAATCGGCAC AAACGGCACG

```

```

351 GCGGATGTTT ACGGGCGCGC TCTCCGTTTC GCCTGTTCTT TGGCAGTCAG
401 GGCGATTTTG TTGCGGACGT AGAGCAAACC GGCGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA AACTTCGCC CATACGAGCG TCCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 134; ORF 0036-1>:

```

m036-1.pep
  1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
 51 VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPLWPMR ESRROSAYPV CLRTAELLPA RTRCLRLKR RIPPAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYER*

```

m036-1/g036 76.8% identity in 228 aa overlap

```

              10      20      30      40      50      60
m036-1.pep  MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAI PWRRHSGAVAIRCSSDSS
              ||||| :||||| ||||| ||||| ||||| ||||| ||||| |||||
g036        MLKPCLVYSACAAALPARTSSSRRCVPSGRCAIYQYSSRADATPRRRHSGAVAIRCSSDSS
              10      20      30      40      50      60

              70      80      90      100     110     120
m036-1.pep  GRFCQTIKAAIPPSFSARKTCS DGETSADSNWRCVHADGLQTASSAASSSQSAQTARMF
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g036        GRFCQTIKAAILPSFSARKTCS DGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF
              70      80      90      100     110     120

              130     140     150     160     170     180
m036-1.pep  TGALSVRPVLWQSGRFCCGRANRRVRHGRQDNRPLWPMRESRRQSAYPVCLRTAELLPA
              : || ||||| ||||| ||| : ||| ||||| ||||| ||| : |||
g036        MFVPSVPPVLWQSGRFCCGRRAVRVRPQRLRDSRRRGRARENRRRSAYRVC LRRADGFPV
              130     140     150     160     170     180

              190     200     210     220     229
m036-1.pep  RTRCLRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPX
              ||| ||||| ||| : |||| | ||||| ||||| ||| ||||| |||
g036        RTHCRCLRKRTPRGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPQRRVCSFAAAAA
              190     200     210     220     230     240

g036        RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX
              250     260     270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 135>:

```

g038.seq
  1 ATGACTGATT TCCGCCAAGA TTCTCTCAA TTCTCCCTCG CCCAAAATGT
 51 TTGGAATTC GGCGAATTTA CCACCAAAGC CGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCCTCTTC AACGACGGCG CGTCCACGCT GCAACTGGCA
151 AAATTCTATG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201 GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC GCGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCAATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGCG CGGTGTCGCC
451 ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTAAATTGT CCGCCGTTCA
501 GGAAGTGGAA AAACAATACG GCCTGCCCGT CGCCCCATC GCCAGCCTGA
551 ACGATTGTG TATCCTGTTG CAAAACAACC CCGAATTGCG ACAGTTCCTC
601 GAACCCGTCC GCACCTACCG CCGGCAGTAC GCGGTAGAAT AA

```

This corresponds to the amino acid sequence <SEQ ID 136; ORF 038.ng>:

```
g038.pep
  1 MTDFRQDFLK FSLAQNVLFK GEFTTKAGRR SPYFFNAGLF NDGASTLQLA
 51 KFYAQSIIES GIRFDMFLGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101 AKDRGEGGVL VGAPLKGRVL IIDVISAGT SVRESIKLIE AEGATPAGVA
151 IALDRMEKGT GKLSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201 EPVRTYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 137>:

```
m038.seq
  1 ATGACCGATT TCCGCCAAGA TTTCTCAAAA TTCTCCCTCG CCCAAAATGT
 51 TTTGAAATTC GGCGAATTCA CCACCAAGGC AGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
151 AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201 GTTCGGTCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC ACGGCGAAGG CGGCGTGTG GTCGGCGCGC CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGCG CGGTGTCGCC
451 ATCGCGCTCG ATCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
501 GGAAGTGGAA AAACAATACG GkCTGCCCCG CGCCCCCATC GCCAGCCTGA
551 ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCCG ACAGTTCCTC
601 GAACCCGTCC GAGCCTACCG TCGGCAGTAC GCGGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 138; ORF 038>:

```
m038.pep
  1 MTDFRQDFLK FSLAQNVLFK GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
 51 KFYAQSIIES GIRFDMFLGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101 AKDHGEGGVL VGAPLKGRVL IIDVISAGT SVRESIKLIE AEGATPAGVA
151 IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201 EPVRAYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 139>:

```
a038.seq
  1 ATGACCGATT TCCGCCAAGA TTTCTCAAAA TTCTCCCTCG CCCAAAATGT
 51 TTTGAAATTC GGCGAATTCA CCACCAAAGC CGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
151 AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201 GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC ACGGCGAAGG CGGCGTGTG GTCGGCGCGC CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGCG CGGTGTCGCC
451 ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
501 GGAAGTGGAA AAACAATACG GCCTGCCCCG CGCCCCCATC GCCAGCCTGA
551 ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCCG ACAGTTCCTC
601 GAACCCGTCC GAGCCTACCG TCGGCAGTAC GCGGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 140; ORF 038.a>:

```
a038.pep
  1 MTDFRQDFLK FSLAQNVLFK GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
 51 KFYAQSIIES GIRFDMFLGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101 AKDHGEGGVL VGAPLKGRVL IIDVISAGT SVRESIKLIE AEGATPAGVA
151 IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201 EPVRAYRRQY GVE*
```

m038/a038 100.0% identity over a 213 aa overlap

```

          10      20      30      40      50      60
m038.pep  MTDFRQDFLKFSLAQNVLFKGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
```

```

a038      |||||
          MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
          10      20      30      40      50      60

          70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
          |||||
a038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
          70      80      90      100     110     120

          130     140     150     160     170     180
m038.pep  IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
          |||||
a038      IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
          130     140     150     160     170     180

          190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
          |||||
a038      ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
          190     200     210

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 038 shows 98.1% identity over a 213 aa overlap with a predicted ORF (ORF 038.ng) from *N. gonorrhoeae*:

m038/g038

```

m038.pep  10      20      30      40      50      60
          MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
          |||||
g038      MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGASTLQLAKFYAQSIIES
          10      20      30      40      50      60

          70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
          |||||
g038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDRGEGGVLVGAPLKGRVL
          70      80      90      100     110     120

          130     140     150     160     170     180
m038.pep  IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
          |||||
g038      IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGKLSAVQEVEKQYGLPVAPI
          130     140     150     160     170     180

          190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
          |||||
g038      ASLNDLFILLQNNPEFGQFLEPVRTYRQYGVEX
          190     200     210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 141>:

g039.seq

```

1   ATGCCGTC CG AACCACCTGC CGCTTCAGAC GGCATCAAAC CGACACACAC
51  CGAGAAAACA TCATGCCCGC CTGTTTCTGT CCGCACTGCA AAACCCGCCT
101 CTGGGTCAA GAAAcccagC TCAAcgtCgC ccaagGCTTC GTCGTCTgcc
151 aaAAAtgcga agGGCTgttt aaAgccaaaG accAtctggc aaGcacGAAA
201 gaacctatat tcaacgattg gcccgaagct gtttcgggat gTcaaaCTCG

```

```

251 TCcaccgcaT cggcacgcac gccattagca aGAaacagat gtccccgcgac
301 gaaatCgccg atatacctcaa cggcggtaca accCTGCACG ATACGCCGCC
351 CGCAACCGCC GCTGCCGCac ctGCCGCCGC ACCGCaggTT TCCGTACCGC
401 CCGCCCGTCA GGAAGGGCTC AACTGGACTA TTGCAACCCT GTTCGCACTT
451 ATCGTCCTCA TTATGCAGCT TTCCTACCTC TTCATCCTAT GA

```

This corresponds to the amino acid sequence <SEQ ID 142; ORF 039.ng>:

**g039.pep**

```

1 MPSEPPAASD GIKPTHTEKT SCPPVSVRTA KPASGSKKPS STSPKASSSA
51 KNAKGCLKPK TIWQARKNLY STIGPKLFRD VKLVHRIGTH AISKKQMSRD
101 EIADILNGGT TLHDTTPATA AAAPAAAPQV SVPPARQEGT NWTIATLFAL
151 IVLIMQLSYL FIL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 143>:

**m039.seq**

```

1 ATGCCGTCCG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51 CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
101 CTGGGTCAA GAAACCCAAC TCAATGTGCG CGnnnnnnnnn nnnnnnnnnn
151 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
201 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnCCC GAGGCTGTTT
251 CGGATGTCAA ACTCGTTCAC CGTATCGGCA CGCGCGCCAT CGGCAAGAAA
301 CAGATTTCCT GTGACGAAAT CGCCGGCATC CTCAACGGCG GTACAACCCA
351 GCCCCGATAT CCGCCCGCAA CCGCCGCCAC CCTGCTGCC GCACCGCAGG
401 TTACCGTACC GCGCGCCGCG CCGCCCGGTC AGGATGGGTT CAACTGGACG
451 ATTGCAACCC TGTTCGCCCT TATCGTCCTC ATTATGCAGC TTTCCTACCT
501 CGTCATCCTA TGA

```

This corresponds to the amino acid sequence <SEQ ID 144; ORF 039>:

**m039.pep**

```

1 MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPPXXXXXX
51 XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX EAVSDVKLVH RIGTRAIGKK
101 QISRDEIAGI LGGTTQPD I PPATAATPAA APQVTVPPAA PARQDGFNWT
151 IATLIFALIVL IMQLSYLVIL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 145>:

**a039.seq**

```

1 ATGCCGTCTG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51 CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
101 CTGGGTCAA GAAACCCAAC TCAATGTGCG CCAAGGCTTC GTCGTCTGCC
151 AAAAATGCGA AGGAATGTTT AAAGCCAAAG ACCATCTGGC AAGCAGGAAA
201 GAACCCATAT TCAACGATT. TGCCCGAAGC TGTTCGGAT GTCAAACCTCG
251 TTCACCGCAT CGGCACGAGC GCCATCGGCA AGAAACAGAT TTCCCGTGAC
301 GAAATCGCCG GCATCCTCAA CGCGGCACA ACCCAGCCCG ATATTCCGCC
351 CGCAACCGCC GCCACCCCTG CTGCCGCACC GCAGGTTACC GTACCGCCCG
401 CCGCGCCCGC CCGTCAGGAT GGGTTCAACT GGACGATTGC AACCTGTTT
451 GCCCTTATCG TCCTCATTAT GCAGCTTTC TACCTCGTCA TCCTATGA

```

This corresponds to the amino acid sequence <SEQ ID 146; ORF 039.a>:

**a039.pep**

```

1 MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPKASSSA
51 KNAKECLKPK TIWQARKNPY STIXPEAVSD VKLVHRIGTS AIGKKQISRD
101 EIAGILNGGT TQPDIPPATA ATPAAAPQVT VPPAAPARQD GFNWTIATLF
151 ALIVLIMQLS YLVIL*

```

**m039/a039** 79.4% identity over a 170 aa overlap

```

          10      20      30      40      50      60
m039.pep  MPSEPPYASDGIKPDTHEEIPCPPVSAPTA KPVSGSKKPN SMSPPXXXXXXXXXXXXXXXXXX
          |||
a039      MPSEPPYASDGIKPDTHEEIPCPPVSAPTA KPVSGSKKPN SMSPKASSSAKNAKECLKPK

```

```

              10      20      30      40      50      60
m039.pep      XXXXXXXXXXXXXXXXXXXXPEAVSDVKLVHRIQTRAIQKKQISRDEIAGILNGGTTQPDI
a039           :      : |      |||||      |||||      |||||      |||||      |||||
              70      80      90      100     110
TIWQARKNPYSTIX-----PEAVSDVKLVHRIQTSAIQKKQISRDEIAGILNGGTTQPDI
              70      80      90      100     110

              130     140     150     160     170
m039.pep      PPATAATPAAAPQVTVPPAAPARQDGFNWTIATLFLALIVLIMQLSYLVILX
a039           |||||      |||||      |||||      |||||      |||||
              120     130     140     150     160
PPATAATPAAAPQVTVPPAAPARQDGFNWTIATLFLALIVLIMQLSYLVILX

```

Computer analysis of this amino acid sequence gave the following results:

### Homology with a predicted ORF from *N. gonorrhoeae*

ORF 039 shows 60.8% identity over a 171 aa overlap with a predicted ORF (ORF 039.ng) from *N. gonorrhoeae*:

m039/g039

```

                                10      20      30      40      50      60
m039.pep  MPSEPPYASDGIKPDTHEEIPC PPVSAPTAKPVSGSKKPNSMSPXXXXXXXXXXXXXXXXXX
          ||||| |||||  |: |||||: ||||:|||||:| ||
g039      MPSEPPAASDGIKPHTHTKTS CPPVSVRTAKPASGSKKPSSTSPKASSSAKNAKGCLKPK
          10      20      30      40      50      60

                                70      80      90      100     110     120
m039.pep  XXXXXXXXXXXXXXXXXXXXPEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPD
          :      :      |: ||||| |||||: |: ||||: ||||| ||||| |
g039      TIWQARKNLYSTIG-----PKLFRDVKLVHRIGTHAISKQMSRDEIADILNGGTTLHDT
          70      80      90      100     110

                                130     140     150     160     170
m039.pep  PPATAAT-PAAAPQVTVPPAAPARQDGFNWTIATLFLALIVLIMQLSYLVILX
          |||||: |||||: |||||: |||||: |||||: |||||: |||||
g039      PPATAAAAPAAAPQVSVPPA---RQEGLNWTIATLFLALIVLIMQLSYLVILX
          120     130     140     150     160

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 147>:

g040.seq

1	ATGAACGCGC	CCGACAGCTT	TGTCGCCAC	TTCCGCGAAG	CCGCCCCCTA
51	CATCCGCCAA	ATGCGCGGCA	CGACACTGGT	CGCCGGCATA	GAcggCCGCC
101	TGCTCGAAGG	CGGCACCTTA	AATAAGCTCG	CCGCCGACAT	CGGGCTGTTG
151	TCGCAactTG	GCATCCGAT	CGTCTCATC	CACGGCGCGT	ACCACTTCCI
201	CGAcgCCTC	CGCGCCGCGC	AAGgccGCAC	GCCGCATTAT	TGCCGgggtt
251	tCGCGTTTAC	CGACGaAACc	tGctctgGAC	AGGCGCAGCA	GtttGCCGGG
301	AccgTCCGCA	GCCGTTTTGA	agcCGCATTG	tgcggcagCG	tttcaggatt
351	cgcgCGCGCG	CCTTCCGTCC	CGCTCGTatc	gggcaacttc	ctgacGCCCC
401	GTCcgatggg	cgtgattgac	ggaACCgata	tggaaatacgc	ggggggttatc
451	cgcaaaaccg	ACACCCGCGC	CCTCGTTTTC	CAACTCGACG	CGGGCAATAT
501	CGTCTGGATG	CCGCGCTCG	GGCATTCTTA	CGGCGGCAAA	ACCTTCAATC
551	TCGATATGGT	GCAGGCGGCC	GCTTCCGTCG	CCGTCTCGCT	TCAGGCCGAA
601	AAACTCGTTT	ACCTGACCCT	TTCAGACGGC	ATTTCCCGCC	CCGACGGCAC
651	GCTCGCCGAA	ACCCTCTCGG	CACAGGAAGC	GCAATCGCTG	CGCGAACACG
701	CCGCGACGCA	AACCCGACGA	CTGATTTCGT	CCGCCGTTGC	CGCGCTCGAA
751	GGCGGCAGTG	ATCGCGTCCA	AATCCTCAAC	GGGGCCGCCG	ACGGCAGCCT
801	GCTGCAAGAA	CTCTTCAACC	GCAACGGCAT	CGGCACGTCC	ATTGCCAAAG
851	AAGCCTTCGT	CTCATCTCCG	CAGGCGCACA	CGGCGCATAC	CCCGCATATC
901	GCCGCCCTCA	TCCGCCCGCT	GGAAGAACAG	GGCGTCTAT	TGCACCGCAG
951	CCGCGAATAC	CTCGAAAACC	ACATTTCCGA	ATTTTCCATC	CTCGAACACG



```

1001 ACGGCGACCT GTACGGCTGT GCCGCACTCA AAACCTTTGC CGAAGCCGAT
1051 TGCGGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGg
1101 ctACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151 GCATAAGCAG GCTGTTCGCA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201 GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGCTGCCCC AAACGCGGCG
1251 CAAAGACTAC CGCAGCAACG GACGAAACCC GCATATTCTG GTGCGTCGCC
1301 TGCACCGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 148; ORF 040.ng>:

**g040.pep**

```

1 MNAPDSFVAH FREAAPYIRQ MRGTTLVAGI DGRLLGGTLL NKLAADIGLL
51 SQLGIRLVLI HGAYHFLDRL AAAQGRTPHY CRGLRVTDET SLGQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPMGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWM PPLGHSYGGK TFNLDMVQAA ASVAVSLQAE
201 KLVYLTSLDG ISRPDGTAE TLSAQEAQSL AEHAASETRR LISSAVAALAE
251 GGVHRVQILN GAADGSLLEQ LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
301 AALIRPLEEQ GVLLHRSREY LENHISEFSI LEHDGDLYGC AALKTFAEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRRKDY RSNGRNPHIL VRRLLHR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 149>:

**m040.seq**

```

1 ATGAGCGCGC CCGACCTCTT TGTCGCCCAC TTCCGCGAAG CCGTCCCCTA
51 CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACCGCC
101 TGCTCGAAGG TGATACCTTA AACAAAGCTCG CCGCCGACAT CCGGCTGTTG
151 TCGCAACTGG GCATCAGGCT CGTCCTCATC CACGGCGCGC GCCACTTCCT
201 CGACCGCCAC GCCGCCGCTC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
251 TGCGCGTTAC CGACGAAACC TCGCTCGAAC AGGCGCAgCA GTTTGCCGGC
301 ACCGTCCGCA GCCGTTTTGA AGCCGCATTG TGCGGCAGCG TTTCCGGGTT
351 CGCGCGCGCG CCTTCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
401 GTCCGATAGG TGTGATTGAC GGAACCGATA TGGAAATACG GGGCGTTATC
451 CGCAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
501 CGTCTGGCTG CCGCGGCTCG GACATTCTTA CAGCGCAAG ACCTTCTATC
551 TCGATATGCT TCAAACCGCC GCCTCCGCCG CCGTCTCGCT TCAGGCCGAA
601 AAATCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
651 GCTCGCCGAA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
701 CCGGCGGGCA AACCGGACGG CTGATTTCGT CCGCCGAACT CTTACCCGC
751 AACGGCATCG GCACGTCCAT TGCCAAAGAA GCCTTCGTCT CCATCCGGCA
801 rGCGCAywgG G.CGACATCC CGCACATCGC CGCCCTCATC CGCCCGCTGG
851 AAGAACAGGG CATCTGCTG CACCGCAs.c GCGAATACCT CGAAAACCAC
901 ATTTCCGAAT TTCCATCCT CGAACACGAC GGCAACCTGT ACGGTTGCGC
951 CGCCCTGAAA ACCTTTGCCG AAGCCGATTG CGGCGAAATC GCCTGCCTTG
1001 CCGTCTCGCC GCag.cACAG GACGGCGGCT ACGGCGAACG CnTGCTTGCC
1051 CACATTATCG ATAAGGCGCG CGGCATAGGC ATAAGCAGGC TGTTCGCACT
1101 GTCCACAAAT ACCGGCGAAT GGTTCGCCA ACGCGGCTTT CAGACGGCAT
1151 CGGAAGACGA GTTGCCCGAA ACGCGGCGCA AAGACTACCG CAGCAACGGA
1201 CGGAACTCGC ATATTCTGGT ACGTCGCCTG CACCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 150; ORF 040>:

**m040.pep**

```

1 MSAPDLFVAH FREAVPYIRQ MRGKTLVAGI DDRLLEGDTL NKLAADIGLL
51 SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFYLDMLQTA ASAAVSLQAE
201 KLVYLTSLDG ISRPDGTAE TLSAQEAQSL AEHAGGQTRR LISSAEFLTR
251 NGIGTSIAKE AFVSIRQAHX XDIPHIAALI RPLEEQGILL HRXREYLENH
301 ISEFSILEHD GNLYGCAALK TFAEADCGEI ACLAVSPQXQ DGGYGERXLA
351 HIIDKARGIG ISRLFALSTN TGEWFAERGF QTASEDELPE TRRKDYRSNG
401 RNSHILVRRL HR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 151>:

**a040.seq**

```

1  ATGATCGTGC CCGACCTCTT TGTCGCCCAC TTCCGCGAAG CCGCCCCCTA
51  CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACCGCC
101 TGCTCGAAGG TGATACCTTA AACAAGTTCG CCGCCGACAT CGGGCTTTTG
151 TCGCAACTGG GCATCAGGCT CGTCCTCATC CACGGCGCGC GCCACTTCCT
201 CGACCGCCAC GCCGCCGCGC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
251 TGCGCGTTAC CGACGAAACC TCGCTCGAAC AGGCGCAGCA GTTTGCCGGC
301 ACCGTCCGCA GCCGTTTGA AGCCGCATTG TGCGGCAGCG TTTCCGGGTT
351 CGCGCGCGCG CCTTCCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
401 GTCCGATAGG TGTGATTGAC GGAACCGATA TGAATACGC GGGCGTTATC
451 CGCAAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
501 CGTCTGGCTG CCGCCGCTCG GACATTCCTA CAGCGGCAAG ACCTTCCATC
551 TCGATATGCT TCAAACCGCC GCCTCCGTCT CCGTCTCGCT TCAGGCCGAA
601 AAACCTCGTT ACCTGACCTT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
651 GCTGCGCGTA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
701 CCGGCGGCGA AACGCGACGG CTGATTTCGT CCGCGGTTGC CGCGCTCGAA
751 GGCGGCGTGC ATCGCGTCCA AATCCTCAAC GGAGCCGCCG ACGGCAGCCT
801 GCTGCAAGAA CTCTTCACCC GCAACGGCAT CGGCACGTCC ATTGCCAAAG
851 AAGCCTTCGT CTCCATCCGG CAGGCGCACA GCGCGGACAT CCCGCACATT
901 GCCGCCCTCA TCCGCCCGCT GGAAGAACAG GGCATCCTGC TGACCCGCAG
951 CCGCAATAC CTCGAAAACC ACATTTCCGA ATTTTCCATC CTCGAACACG
1001 ACGGCAACCT GTACGGTTGC GCCGCCCTGA AAACCTTTGC CGAAGCCGAT
1051 TGCGGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGG
1101 CTACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151 GCATAAGCAG GCTGTTTCGA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201 GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGTTGCCCG AAACGCGGCG
1251 CAAAGACTAC CGCAGCAACG GACGGAAGTC GCATATTCTG GTGCGTCGCC
1301 TGCACCGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 152; ORF 040.a>:

```

a040.pep
1  MIVPDLFVAH FREAAPYIRO MRGKTLVAGI DDRLLEGDTL NKFAADIGLL
51  SOLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFHLDMLQTA ASVAVSLQAE
201 KLVYLTSDG ISRPDGTAVL TSAQEAQSL AEHAGGETRR LISSAVAAL
251 GGVHRVQILN GAADGSLLQE LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
301 AALIRPLEEQ GILLHRSREY LENHISEFSI LEHDGNLYGC AALKTFEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRRKDY RSNGRNSHIL VRRLHR*

```

m040/a040 91.5% identity in 436 aa overlap

```

          10      20      30      40      50      60
m040.pep  MSAPDLFVAHFREAVPYIROMRGKTLVAGIDDRLLEGDTL NKLAADIGLLSOLGIRLVLI
          | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a040       MIVPDLFVAHFREAAPYIROMRGKTLVAGIDDRLLEGDTL NKFAADIGLLSOLGIRLVLI
          10      20      30      40      50      60

          70      80      90     100     110     120
m040.pep  HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a040       HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA
          70      80      90     100     110     120

          130     140     150     160     170     180
m040.pep  PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a040       PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK
          130     140     150     160     170     180

          190     200     210     220     230     240
m040.pep  TFYLDMLQTAASAAVSLQAEKLVYLTSLDGISRPDGT LAETLSAQEAQSLAEHAGGQTRR
          | : | | | | | | | | | | | | | | | | | | | | | | | | | |
a040       TFHLDMLQTAASVAVSLQAEKLVYLTSLDGISRPDGT LAETLSAQEAQSLAEHAGGETRR

```

	190	200	210	220	230	240
m040.pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI					
a040	LISSAVAALGGVHRVQILNGAADGSLLOELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI					
	250	260	270	280	290	300
m040.pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS					
a040	AALIRPLEEQGILLHRSREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS					
	310	320	330	340	350	360
m040.pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY					
a040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY					
	370	380	390	400	410	420
m040.pep	RSNGRNSHILVRRLHRX					
a040	RSNGRNSHILVRRLHRX					
	430					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 040 shows 88.3% identity over a 436 aa overlap with a predicted ORF (ORF 040.ng) from *N. gonorrhoeae*:

m040/g040

m040.pep	MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI	60
g040	MNAPDSFVAHFREAAPYIRQMRGTTLVAGIDGRLLGGTLNKLAADIGLLSQLGIRLVLI	60
m040.pep	HGARHFLDRHAAAQGRTPHYCRGLRVTDSETSLEQAQQFAGTVRSRFEAALCGSVSGFARA	120
g040	HGAYHFLDRLAAAQGRTPHYCRGLRVTDSETSLGQAQQFAGTVRSRFEAALCGSVSGFARA	120
m040.pep	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK	180
g040	PSVPLVSGNFLTARPMGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWMPPLGHSYGGK	180
m040.pep	TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRDPDGTLAETLSAQEAQSLAEHAGGQTRR	240
g040	TFNLDMVQAAASVAVSLQAEKLVYLTLSDGISRDPDGTLAETLSAQEAQSLAEHAASETRR	240
m040.pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI	276
g040	LISSAVAALGGVHRVQILNGAADGSLLOELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI	300
m040.pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS	336
g040	AALIRPLEEQGVLLHRSREYLENHISEFSILEHDGDLYGCAALKTFAEADCGEIACLAVS	360
m040.pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY	396
g040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY	420
m040.pep	RSNGRNSHILVRRLHRX	413
g040	RSNGRNPILVRRLHRX	437

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 153>:

```
g041.seq
1   ATGAGTTCGC CCAAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGCCT
51  GATTACCGCC GCCGCCTTCG TGC GCGAACC GCAAAGCATC GGTGCGCTGG
101 TGTGCGAAGT ACCGCTGACC GATATGATCC GTTATCCGCT GCTGTCCGCC
151 GGTTC AAGTT GGACGGACGA ATACGGCAAT CCGCAGAAAT ACGAAGCCTG
201 CAAACGCCGG CTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCACTC ATTACCACCA GCCTCAGCGA CGACCGCGTC
301 CATCCCGCCC ACGCGCTCAA ATTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCCAACGCGA ATCCGCCGAC AAACCTGCCT GCGTGTGCT GTTTTTGAAA
451 GAATTTTGG GATAA
```

This corresponds to the amino acid sequence <SEQ ID 154; ORF 041.ng>:

```
g041.pep
1   MSSPKHIGLQ GGSNGGLITA AAFVREPOSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQKYEACKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPOSW LYSPDGGGHT GNGTQRESAD KLACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 155>:

```
m041.seq
1   ATCAGTTCGC CCGAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGACT
51  GATTACTGCC GCCGCCTTCG TGC GCGAACC GCAAAGCATC GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTC AAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCCAACGCGA ATCCGCCGAC GAACTCGCCT GCGTCTTGCT GTTTTTGAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 156; ORF 041>:

```
m041.pep
1   ISSPEHIGLQ GGSNGGLITA AAFVREPOSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQKYEVCRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSQSW LYSPDGGGHT GNGTQRESAD ELACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 157>:

```
a041.seq
1   ATCAGTTCGC CCGAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGACT
51  GATTACTGCC GCCGCCTTCG TGC GCGAACC GCAAAGCATA GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTC AAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CGCAGCGCGA AGCCGCCGAC GAACTCGCCT GCGTGTGCT GTTTTTGAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 158; ORF 041.a>:

```
a041.pep
1   ISSPEHIGLQ GGSNGGLITA AAFVREPOSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQKYEVCRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPOSW LYSPDGGGHT GNGTQREAAD ELACVLLFLK
```

151 EFLG\*

m041/a041 98.7% identity over a 154 aa overlap

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
a041	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRHPAHALKFYAKLRETSAQSW					
a041	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRHPAHALKFYAKLRETSQPSW					
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX					
a041	LYSPDGGGHTGNGTQREAADELACVLLFLKEFLGX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 041 shows 96.8% identity over a 154 aa overlap with a predicted ORF (ORF 041.ng) from *N. gonorrhoeae*:

m041/g041

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	:					
g041	MSSPKHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRHPAHALKFYAKLRETSAQSW					
g041	PQKYEACKRRLGELSPYHNLSGDIDYPPALITTSLSDDRHPAHALKFYAKLRETSQPSW					
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX					
g041	LYSPDGGGHTGNGTQRESADKLACVLLFLKEFLGX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 159>:

g041-1.seq

```

1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAAACAA CGACAAGGCG CGCGCACTTT CAGACGGCAT TTTGAATCAA
151 ATGCAGGACA CGCGGCAGAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTT CATCAGAATG CGGAATATCC GAAGGGCGTG TACCGCATGT
251 GTACGGCGGC GACCTACCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTTCGATG GTTGCTCGGC GACGATGTGT ATTTGGGCGG
351 CGTGTCGCAC TTGGTGGAGC AGCCCAACCG CGCGCTGCTG ACTTTGAACA
401 AATCGGGCGG CGATACGGCG TATACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAATTGGTAG AGGGCGGTTT TCACTTTCCG GCAGGCAGAA ACCATGTGTC
501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
551 AGTTGACCGA ATCGGGCTAT CCGCGCGAAG TGTGGCTGGT GGAACGCGGC

```

```

601 AAGAGTTTCG AGGAAAGCCT GCCGGCGTAC CAAATCGATA AAGGCGCGAT
651 GATGGTAAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCGGACGGT TTTTACACCA AGACGTATTT GCAGGTGTCTG
751 TCCGAAGGCG GGGCGAAACC GTTGAACCTG CCTAATGATT GCGATGTGGT
801 CCGCTATCTG GCGGGACATC TTTTGCTGAC GCTGCGCAAG GAACTGAAT
851 GCGCGAACCA AAGCTATCCG AGTGGCGCGT TGGTGGCGGT GAACTGAAT
901 CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
951 GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCAAGCCTGC
1001 TGGAGAATGT ACAAGGCCGT CTGAAAGCGT GGCGGTTTGC CGACAGCAAA
1051 TGGCAGGAAG CCGAGTTGCC GCACCTGCCC TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGCGGCG ACGTGGTTTA TCTTGCCGCC AGCGATTTCA
1151 CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAAGTACC
1201 GTCATGCGCC TCCAGCCGCA GCAGTTTGTT TCAGACGGCA TCGAAGTGCG
1251 GCAGTTTGG GCGGTGTCTG CCGACGGCGA ACGCATTCTT TATTTCCACG
1301 TCGGCAAAAA CGCCGCGCCC GACACGCCGA CCTTAGTCTA TGCTTACGGA
1351 GGTTCGGCA TTCCTGAATT GCCGCATTAT CTGGGCAGCG TCGGCAATA
1401 TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCAACATC CGCGGCGGCG
1451 GAGAATTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAC
1501 AAAAGCGTTG ATGATTGTG GGCAGTCGTG CGTGATTGTG CCGAACGCGG
1551 CATGAGTTCG CCCAAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGCC
1601 TGATTACCGC CGCCGCTTTC GTGCGCGAAC CGCAAAGCAT CGGTGCCGTG
1651 GTGTGCGAAG TACCGCTGAC CGATATGATC CGTTATCCGC TGCTGTCCGC
1701 CGGTTCAAGT TGGACGGACG AATACGGCAA TCCGCAGAAA TACGAAGCCT
1751 GCAAACGCCG GCTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCGCACT CATTACCACC AGCCTCAGCG ACGACCGCGT
1851 CCATCCGCGC CACGCGCTCA AATTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACCCAACGCG AATCCGCCGA CAAACTCGCC TCGTGTTGCG TGTTTTTGAA
2001 AGAATTTTGG GGATAA

```

This corresponds to the amino acid sequence <SEQ ID 160; ORF 041-1.ng>:

g041-1.pep

```

1 MKSYDPYRHF FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILNQ
51 MQDTRQIPFC QEHRARMYHF HQNAEYPKGV YRMCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLNKSGGDTA YTLEVDLEAG
151 ELVEGGFHFAP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPAY QIDKGAMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251 SEGGAKPLNL PNDQDVVGYL AGHLLLTLRK DWHRANQSYR SGALVAVKLN
301 RGELGAAQLL FAPDETQALE SVETTKRFV ASLLENVQGR LKAWRFADSK
351 WQEAELPHLP SGALMTDQF WGGDVVYLAA SDETTPLTLF ALDLNVMELT
401 VMRLQPQQFV SDGIEVRQFW AVSSDGERIP YFHVGNKNAAP DTPTLVYAYG
451 GFGIPELPHY LGSVGYWLE EGNFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLLAVV RDLSEGMSS PKHIGLQGS NGGLITAAAF VREPQSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEACKRRLGE LSPYHNLSHG
601 IDYPPALITT SLSDRVHFA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQRESADKLA CVLLFLKEFL G*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 161>:

m041-1.seq

```

1 ATGAAATCCT ACCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51 CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAGAAAA CGACAAGGCG CGCGCGCTTT CAGACGGCAT TTTGGCGCAG
151 TTGCAGGACA CGCGGCAGAT TCCGTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTT CATCAGGACG CGGAGTATCC GAAGGGCGTG TACCGCGTGT
251 GTACCGCGGC GACGTATCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTTCGACGA ATTGCTTGGC GACGATGTGT ATTTGGCGCG
351 CGTGTGCGAC TTGGTGGAA AGCCCAACCG CGCGTTGTGA AACTGAGCA
401 AATTGGGCAG CGATACGGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAGTTGGTCG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG AACGAACGCC
551 AGTTGACCCA ATCGGGCTAT CCGCGCGAAG TATGGCTGGT GGAACGCGGC
601 AAGAGTTTCG AGGAAAGCCT GCCTGTGTAT CAAATCGGCG AAGACGGCAT
651 GATGGTGAAC GCGTGGCGTT ATCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCGGACGGT TTTTACACCA AAACCTATTT GCGGGTCTCA

```

m041-1.pap

1	MKSYDPDPYRH	FENLDSAETQ	NFAAEANAET	RARFLENDKA	RALSDGILAQ
51	LQDTRQIPFC	QEHRARMYHF	HQDAEYPKGV	YRVCTAATYR	SGYPEWKILF
101	SLVDFDELLG	DDVYLGGVSH	LVEQPNRALL	TLSKLQSDTA	YTLVEVDLEAG
151	EDVEGGFHPF	AGKNHVSWRD	ENSVWCPRAW	NERLGTOSGY	PREVWLVERG
201	KSFEESLPVY	QIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG	FYTKTYLRVS
251	AEGEAGPLNL	PNDCDVVGYL	AGHLLTLRLK	DWNRANQSYSP	SGALVAVKLN
301	RGELGAQAQL	FAPDETQALE	SVETTKRFVV	ASLLENVQGR	LKAWRFADKG
351	WQEVELPRLP	SGALEMTDQP	WGGDVVYLAA	SDFTTPLTLF	ALDLNVMELT
401	VMRRQPOQFD	SDGINQQWF	TTSADGERIP	YFHWGKNAAP	DMPTLVYAYG
451	GFGIPELPHY	LGSIGKYWLE	EGNAFVLNRI	RGGGEFGPRW	HQAQAQISIKH
501	KSVDDLLAVV	RDLSEIRGISS	PEHIGLQGGG	NGGLITAAAF	VREPOSIGAL
551	VCEVPLTDMI	RYPLLSAGSS	WTDEYGNPQK	YEVCKRRLGE	LSPYHNLSDG
601	IDYPPALITT	SLSDRVHPA	HALKFYAKLR	ETSAQSWLYS	PDGGGHTNGG
651	TORESADELA	CVLLFLKEFL	G*		

**m041-1/g041-1** 94.6% identity in 671 aa overlap

	10	20	30	40	50	60
m041-1.pep	MKSYPDPYRHFENLDSAETQNF	AAEANAETRARFL	ENDKARALSDG	ILAQ	LQDTRQ	IPFC
g041-1	MKSYPDPYRHFENLDSAETQNF	AAEANAETRARFL	NNDKARALSDG	ILNQ	MQDTRQ	IPFC
	10	20	30	40	50	60
	70	80	90	100	110	120
m041-1.pep	QEHRARMYHFHQDAEY	PKGVYRVCTAATY	RSGYPEWKILFS	VADFDELLGDD	VYLG	GVSH
g041-1	QEHRARMYHFHQDAEY	PKGVYRMCTAATY	RSGYPEWKILFS	VADFDELLGDD	VYLG	GVSH
	70	80	90	100	110	120
	130	140	150	160	170	180
m041-1.pep	LVEQPNRALLTSLK	LGSDTAYTLEVD	LEAGELVEGGF	HFPAGKNHVS	WRDENS	VWVCPAW
g041-1	LVEQPNRALLTSLK	SGDGTAYTLEVD	LEAGELVEGGF	HFPAGKNHVS	WRDENS	VWVCPAW
	130	140	150	160	170	180
	190	200	210	220	230	240
m041-1.pep	NERQLTQSGYPRE	VWLVERGKSFEES	LPVYQIGEDGM	MVNAWRYLDP	QGSPI	DLIEASDG
	:	:	:	:	:	:

Query: 121 LVEQPNRALLTSLKLGSDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVMVWCPAW 180  
L N L+ LS G D E D+ GE V+ GF P GK V+W DEN+++V W



Sbjct: 162 LPPTSNLCLIRLSDGGKDADVREFDIAKGEFVKEGFVLPEGKQSVTWVDENTIIYVTREW 221

Query: 181 NERQLTQSGYPREVWLVERGKSFEESLPVYQ-----IGEDGMM--VNAWRYLDPQGSPI 232  
 ++T SGY +V+RG+S ++++ +++ E G++ ++ +D +

Sbjct: 222 TPGEVTSSTGYAYVTKVVKRGQSLDQAVEIFRGQKQKDVSAERGVLRDIDGKYVMDTSYRGL 281

Query: 233 DLIEASDGFYTKTYLRVSAEAGEAKPLNLPNDPNDVGVYLAGHLLTLRKDWNANQS-YPS 291  
 D FY + + L LP GY G + L+ DW A + + +

Sbjct: 282 DFFNTELAIFYPNHG---PDTRKVVLPLPTTAVFSGYKQAIYWLKSDWTSAGKTVFHN 337

Query: 292 GALVAVKLNIRGELGAAQL---LFAPDETOALESVETTKRFVVASLLENVQGRKAWRFA 347  
 GA++A L A++ LF P+E Q++ TK +V S+L NV ++++ F

Sbjct: 338 GAIIAFDLKAALADPARVEPLVLFMPNEHQSVAGTTQTKNRLVLSILSNVTSEVRSFDFG 397

Query: 348 DGKWQEVELPRLPSGALEMTDQPWGGDVVYLAASDFTTPTLTLFALDLNVMELTMVRRQFQ 407  
 G W +L + L +T D +++ + F P TLF D ++ + P

Sbjct: 398 KGGWSSFKLALPENSTLSLTSSDDESQDLFVFSEGFLEPSTLFCADAATGQVEKITSTPA 457

Query: 408 QFDSGGINVQQFWTTSADGERIPYFHVGNKNAAP---DMPTLVYAYGGFGIPELPHYLGSI 464  
 +FD+ G+ QQFW TS DG ++PYF V + PT++YAYGGF IP P Y +

Sbjct: 458 RFDAGGLQAQFWATSKDGTQVYPFLVARKDVKLDGNTPTILYAYGGFQIPMQPSYSAVL 517

Query: 465 GKYWLEEGNAFVLANIRGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSEGGISSPEHI 524  
 GK WLE+G A+ LANIRGGGEFGP+WH A ++ + DD AV +DL + ++S H+

Sbjct: 518 GKLWLEKGGAYALANIRGGGEFGPKWH DAGLKTNRQRVYDDFQAVAQDLIAKKVTSTPHL 577

Query: 525 GLQGGSSNGGLITAAAFVREPOSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVC 584  
 G+ GGSSNGGL+ ++ P A+V +VPL DM+ + +SAG+SW EYG+P V

Sbjct: 578 GIMGSSNGGLMGVQMIQRPDLWNAVVIQVPLLDVMNFTMSAGASWQAEYGPDD-PVE 636

Query: 585 KRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPFAHALKFYAKLRETSASQSWLYSPDGG 644  
 L +SPYHN+ G+ YP TS DDRV P HA K A + + Y G

Sbjct: 637 GAFLRSISPYHNKAGVAYPEPFETSTKDDRVPVHARKMAALFEDMGLPFYYYENIEG 696

Query: 645 GHTGNGTQRESADELACVLLFL 666  
 GH +E A A +++

Sbjct: 697 GHAAAAANLQEHARRYALEYIYM 718

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 163>:

a041-1.seq

```

1  ATGAAATCCT ACCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101  TTTTAAACAA CGACAAGGCA CGCGCATTGT CTGACGGCAT TTTGGCGCAG
151  TTGCAGGACA CGCGGCAAAT TCCGTTTGT CAGGAACACC GCGCGCGGAT
201  GTACCATTTC CATCAAGATG CGGAATATCC GAAAGGCGTG TACCGCGTGT
251  GTACCGCGGC GACTTACCGT TCGGGCTATC CTGAGTGGAA AATCCTGTTT
301  TCGGTGGCGG ATTTGACGCA ATTGCTCGGT GACGATGTAT ATCTAGGCGG
351  CGTGTGCGAC CTGGTGGAAC AGCCCAACCG CGCGTTGTTA AACTGAGCA
401  AATCGGGCGG CGATACCGCG TACACGCTGG AAGTGGATT GGAAGCAGGG
451  GAGTTGGTAG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501  GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
551  AGTTGACCGA ATCGGGCTAT CCGCGCAGG TGTGGCTGGT GGAACGCGGC
601  AAGAGTTTCG AGGAAAGCCT GCCGGTGTAC CAAATTGCTG AAGACGGCAT
651  GATGGTGAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCC CCGATTGATT
701  TGATTGAAGC GTCTGACGGT TTTTACACCA AAACCTATTT GCAGGTCTCA
751  GCCGAAGGCG AAGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTAGT
801  CGGCTATCTG GCCGGACATC TTTTGTGAC CTTGCGTAAA GACTGGCACC
851  GCGCGAACCA AAGCTATCCG AGTGGCGCAT TGGTAGCAGT AAAATTAAAC
901  CGCGGCGAAT TGGGCGCGGC GCAGCTTTTG TTTGCGCCCA ATGAAACGCA
951  GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTCTGT GCGAGCCTGC
1001  TGGAAAACGT ACAGGGTCGT CTGAAAGCGT GGCGTTTAC TGATGGCAAA
1051  TGGCAGGAAA CCGAGTTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101  CGACCAACCG TGGGGGGGCG ACGTAGTTTA CCTTGCCGCC AGCGATTTC
1151  CCACGCCGCT GACGCTGTTT GCATTGGATT TGAACGTGAT GGAACGTACC
1201  GTCATGCGCC GCCAGCCGCA GCAGTTTGAT TCAGACGGCA TTAACGTGCA
1251  GCAGTTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCCACG
1301  TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGCTGGTCTA TGCCACGGC
1351  GGTTCGGCA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAATA
1401  TTGGCTGGAA GAGGGCAATG CCTTGTGATT GGCGAACATC CGCGCGCGGC

```

```
1451 GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501 AAAAGCGTTG ATGATTTATT GGCAGTCGTG AGCGATTTGT CCGAACGCGG
1551 TATCAGTTCG CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601 TGATTACTGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT AGGCGCGCTG
1651 GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701 CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751 GCAAACGCCG GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851 CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACGCAGCGCG AAGCCGCCGA CGAACTCGCC TCGTGTTCG TGTTTTTGAA
2001 AGAGTTTTTG GGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 164; ORF 041-1.a>:

a041-1.pep

```
1 MKSYDPYRH FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILAQ
51 LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TSKSGGDTA YTEVDLEAG
151 ELVEGGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPVY QIAEDGMMVN AWRYLDPOGS PIDLIEASDG FYTKTYLQVS
251 AEGEAKPLNL PNDCDVVGYL AGHLLTLRKH DWHRANQSYR SGALVAVKLN
301 RGELGAAQLL FAPNETQALE SVETTKRFVV ASLLENVQGR LKAWRFTDGK
351 WQETELPRLP SGALEMTDQP WGGDVVYLAA SDFTTPTLTF ALDLNVMELT
401 VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGNKNAAP DMPTLVYAYG
451 GFGIPELPHY LGSIGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLAVV SDLSERGISS PEHIGLQGS NGGLITAAAF VREPOSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQREAADELA CVLLFLKEFL G*
```

a041-1/m041-1 97.9% identity in 671 aa overlap

```
10 20 30 40 50 60
a041-1.pep MKSYDPYRHFENLDSAETQNFNFAAEANAETRARFLNNDKARALSDGILAQ LQDTRQIPFC
m041-1 MKSYDPYRHFENLDSAETQNFNFAAEANAETRARFLNNDKARALSDGILAQ LQDTRQIPFC
10 20 30 40 50 60
70 80 90 100 110 120
a041-1.pep QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH
m041-1 QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH
70 80 90 100 110 120
130 140 150 160 170 180
a041-1.pep LVEQPNRALLT LSKSGGDTAYTLEVDLEAGELVEGGGFHFPAGKNHVSWRDENS VVWVCPAW
m041-1 LVEQPNRALLT LSKSGGDTAYTLEVDLEAGELVEGGGFHFPAGKNHVSWRDENS VVWVCPAW
130 140 150 160 170 180
190 200 210 220 230 240
a041-1.pep DERQLTESGYPREVWLVERGKSFEESLPVYQIAEDGMMVN AWRYLDPOGS PIDLIEASDG
m041-1 DERQLTESGYPREVWLVERGKSFEESLPVYQIAEDGMMVN AWRYLDPOGS PIDLIEASDG
190 200 210 220 230 240
250 260 270 280 290 300
a041-1.pep FYTKTYLQVSAEAEAKPLNL PNDCDVVGYL AGHLLTLRKH DWHRANQSYR SGALVAVKLN
m041-1 FYTKTYLQVSAEAEAKPLNL PNDCDVVGYL AGHLLTLRKH DWHRANQSYR SGALVAVKLN
250 260 270 280 290 300
310 320 330 340 350 360
a041-1.pep RGELGAAQLL FAPNETQALE SVETTKRFVVASLLENVQGR LKAWRFTDGK WQETELPRLP
m041-1 RGELGAAQLL FAPNETQALE SVETTKRFVVASLLENVQGR LKAWRFTDGK WQETELPRLP
310 320 330 340 350 360
```

	370	380	390	400	410	420
a041-1.pep	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTMRRQPQQFSDGINVQQFW					
m041-1	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTMRRQPQQFSDGINVQQFW					
	370	380	390	400	410	420
	430	440	450	460	470	480
a041-1.pep	TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI					
m041-1	TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLANI					
	430	440	450	460	470	480
	490	500	510	520	530	540
a041-1.pep	RGGGEFGPRWHQAAQGISKHKSVDLLAVVSDLSESGISSPEHIGLQGGSNGLITAAAF					
m041-1	RGGGEFGPRWHQAAQGISKHKSVDLLAVVSDLSESGISSPEHIGLQGGSNGLITAAAF					
	490	500	510	520	530	540
	550	560	570	580	590	600
a041-1.pep	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
m041-1	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
	550	560	570	580	590	600
	610	620	630	640	650	660
a041-1.pep	IDYPPALITTSLSDDRVPALHAKFYAKLRETSQSWLYSPDGGGHTGNGTQREAADELA					
m041-1	IDYPPALITTSLSDDRVPALHAKFYAKLRETSQSWLYSPDGGGHTGNGTQRESADELA					
	610	620	630	640	650	660
	670					
a041-1.pep	CVLLFLKEFLGX					
m041-1	CVLLFLKEFLGX					
	670					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 165>:

**g042.seq**

```

1  ATGACGATGA TTTGCTTGCG CTTCCAagcG TTCGTGCCGC ATACCAGCGC
51  GTTATCCAAC ACTTCCACGG CAGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TGC GGTCGAT GATGAAAATC CAGCCGGGGT TTTTCTCTTT GATGTATTCT
151 AAGGAAACGG GCTGCCCGTG CCCTTCGTTG CGTAAAGATT CGTCCACGGG
201 CGGCAGGCCG ATGTCGCCGT GTATCCAAC TGC CAACCGC GATTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCTGTAACCG ACAGCACCAG CCCGCGTCTT
301 TTGCCTTTGG cggCTTCGCG CTTTGGGCG AACAGCGCGT CAATCTGCGC
351 ATTCAATTCC GCCACGCGCG CTTCTTACC GAAAATCCGC GACAGGGTCT
401 CCATCTGCTT CTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAAA
451 TCTATGgtgG TCGCGTTTTT CGCCAACGTG TCATACGCTT CCGCACCCGG
501 CCCGCCGGTA ATGACAAACT GCGGATTGTG GCGGTGCAGG GATTGCGAAT
551 CGGGCTCAAA CAGCGTCCCC ACCGTTGCCG CTTGTGCAAA TGCAGGCTGC
601 AAATAG

```

This corresponds to the amino acid sequence <SEQ ID 166; ORF 042.ng>:

**g042.pep**

```

1  MTMICLRFOA FVPHTSALSN TSTAAGPSCP MAAVRSMMKI QPGFFSLMYS
51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVDSTSPRP
101 LPLAASRFWA NSASICAFNS ATRASLPKIR DRVSICFSPV VRILPLSTVK
151 SMVVAFFANC SYASAPGPPV MTNCGLWRCR DSQSGSNSVP TVAALSNAGC
201 K*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 167>:

## m042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAmT ACTTCGACAG CCGcCGGCCy TTCyTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTTCG
151 AAGGAAACAG GCTGCCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCGCTG GTGCGGATAT TGCCGTGTGC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTGCGG CTTGTCAAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 168; ORF 042>:

## m042.pep

```

1 MTMICLRFOA FVPRTSALSX TSTAAGXSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPGPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVDSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTSXGLXRCR ASXSGSNSVP TVAALSNAGC
201 K*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 169>:

## a042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCTGCCCC ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTTCG
151 AAGGAAACAG GCTGCCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCGCTG GTGCGGATAT TGCCGTGTGC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCCAACGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTGCGG CTTGTCAAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 170; ORF 042.a>:

## a042.pep

```

1 MTMICLRFOA FVPRTSALSX TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPGPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVDSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*GL*RCR AS*SGSNSVP TVAALSNAGC
201 K*

```

m042/a042 99.0% identity over a 201 aa overlap

	10	20	30	40	50	60
m042.pep	MTMICLRFOAFVPRTSALSXTSTAAGXSCPMMAVRSMMKIQSGFFSLMYSKETGCPGPSL					
a042	MTMICLRFOAFVPRTSALSNTSTAAGPSCPMMAVRSMMKIQSGFFSLMYSKETGCPGPSL					
	10	20	30	40	50	60
m042.pep	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRVWANSASICAFNS					
a042	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
m042.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
a042	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
	130	140	150	160	170	180
m042.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
a042	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					

```

a042      |||||
          AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR
          130      140      150      160      170      180

          190      200
m042.pep  ASXSGSNSVPTVAALSNAGCKX
          |||||
a042      ASXSGSNSVPTVAALSNAGCKX
          190      200

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 042 shows 93.0% identity over a 201 aa overlap with a predicted ORF (ORF 042.ng) from *N. gonorrhoeae*:

m042/g042

```

          10      20      30      40      50      60
m042.pep  MTMICLRFAQFVPRTSALSXTSTAAGXSCPMMAVRSMMKIQSGFFSLMYSKETGCPCPSL
          |||||:|||||
g042      MTMICLRFAQFVPHTSALSNTSTAAGPSCPMMAVRSMMKIQPGFFSLMYSKETGCPCPSL
          10      20      30      40      50      60

          70      80      90      100     110     120
m042.pep  RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
          |||||:|||||
g042      RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRFWANSASICAFNS
          70      80      90      100     110     120

          130     140     150     160     170     180
m042.pep  AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR
          |:|||||:|||||:|||||:|||||: || |||
g042      ATRASLPKIRDVSICFSPLVRILPLSTVKSMVVAFFANC SYASAPGPPVMTNCGLWRCR
          130     140     150     160     170     180

          190     200
m042.pep  ASXSGSNSVPTVAALSNAGCKX
          | |||||
g042      DSQSGSNSVPTVAALSNAGCKX
          190     200

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 171>:

m042-1.seq

```

1   ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51  GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACAG GCTGCCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTGTGC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA A

```

This corresponds to the amino acid sequence <SEQ ID 172; ORF 042-1>:

m042-1.pep

```

1   MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51  KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*

```

m042-1/g042 95.4% identity in 173 aa overlap

240

```

              10      20      30      40      50      60
m042-1.pep  MTMICLRFOAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPGPSL
              10      20      30      40      50      60
g042         MTMICLRFOAFVPHTSALSNTSTAAGPSCPMAAVRSMMKIQPGFFSLMYSKETGCPGPSL
              10      20      30      40      50      60
m042-1.pep  RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS
              10      20      30      40      50      60
g042         RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS
              10      20      30      40      50      60
m042-1.pep  AARASLPKIRAKVSICFSPVLRILPLSTVRSMVVAFFANC SYASAPGPPVMTSX
              10      20      30      40      50      60
g042         ATRASLPKIRDRVSICFSPVLRILPLSTVKSMVVAFFANC SYASAPGPPVMTNCGLRWRCR
              10      20      30      40      50      60
g042         DSQSGSNSVPTVAALSNAGCKX
              10      20      30      40      50      60

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 173>:

```

a042-1.seq
1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGCGCG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCG
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCTT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCCAACGTG TCATACGCTT CCGCGCCCGG
501 CCCGCCGTA A

```

This corresponds to the amino acid sequence <SEQ ID 174; ORF 042-1.a>:

```

a042-1.pep
1  MTMICLRFOA FVPRTSALS N TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPGPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVT DSTSPR
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPV VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*

```

m042-1/a042-1 100.0% identity in 173 aa overlap

```

              10      20      30      40      50      60
m042-1.pep  MTMICLRFOAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPGPSL
              10      20      30      40      50      60
a042-1       MTMICLRFOAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPGPSL
              10      20      30      40      50      60
m042-1.pep  RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS
              10      20      30      40      50      60
a042-1       RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS
              10      20      30      40      50      60
m042-1.pep  AARASLPKIRAKVSICFSPVLRILPLSTVRSMVVAFFANC SYASAPGPPVMTSX
              10      20      30      40      50      60
a042-1       AARASLPKIRAKVSICFSPVLRILPLSTVRSMVVAFFANC SYASAPGPPVMTSX
              10      20      30      40      50      60

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 175>:

```

g043.seq
  1 ATGGTTGTTT CAAATCAAAA TATCTATGCC GTCGGCCCAT CAGCACTTTT
 51 TCACATCCGA AGGCAAAAAT CCGTAATGCC GCCTGAACGC TTCgttgaAC
101 CGTCCCGCGT ggcggtagcc gcAAAAGTGC ATcGCGGCTT GGATGGTGCT
151 GCCCGATTTCG ATGAGGGcga gcGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201 GTCCGGCGAC GGTTCGCGG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTCGTTCA GCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGGCG
301 GGCGAATTTCG CTGTTCAAAA TATCGGCGGC TTCGTCTATG CGCCGGCGGC
351 GGTAGCCGTT GTCGTGGCGG CGGAAGGTGA AGCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 176; ORF 043.ng>:

```

g043.pep
  1 MVVSNQNIYA VGPSALFHIR RQKSVMPPER FVEPSRVAVA AKVHRGLDGA
 51 ARFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQPD A GDFGDGQRA
101 GEFAVQNIGG FVYAPAAVAV VVAAEGEA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 177>:

```

m043.seq
  1 ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCCT CAGCACTTCT
 51 TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101 CGTCCCGCGT GCGGGTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT
151 GCCGGATTTCG ATGAGGGCGA GCGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201 ATCCGGCGAC GGTTCGCGG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTCGTTCA GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
301 GGCGAATTTCG TGTTGCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
351 GTTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 178; ORF 043>:

```

m043.pep
  1 MVVSNQNIYA AGPSALLHIR RQKSVMPSEF FVEPSRVAVA AKVHGGLDGA
 51 AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*

```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis* *menA* with *menB*

ORF 043 shows 89.8% identity over a 128 aa overlap with a predicted ORF (ORF043.a) from *N. gonorrhoeae*:

m043/g043

	10	20	30	40	50	60
m043.pep	MVVSNQNIYAAGPSALLHIRRQKSVMPSEFVEPSRVAVAAKVHGGLDGAAGFDEGERVF					
	:    :					
g043	MVVSNQNIYAVGPSALFHIRRQKSVMPPERFVEPSRVAVAAKVHRGLDGAARFDEGERVF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m043.pep	QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV					
	:    :                         : : : : : : : :					
g043	QPQAAQASGDGFAGLRFEIAFQVAFVQPDAAAGDFGDGQRTAGEFAVQNIGGFVYAPAAVAV					
	70	80	90	100	110	120
	130					
m043.pep	VVAAEGEAQX					
g043	VVAAEGEAXX					
	130					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 179>:

```

a043.seq
  1 ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCCT CAGCACTTCT
 51 TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101 CGTCCCGCGT GCGGGTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT

```

```

151 GCCGGATTTC ATGAGGGCGA GCGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201 ATCCGGCGAC GGTTCGCGG GTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTCGTTCA GTCCGACGCG GCGGGCGATT TCGCGGATGG TCAGCGGACG
301 GGCGAATTCG TGTTCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
351 GGTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 180; ORF 043.a>:

```

a043.pep
1  MVVSNQNIYA AGPSALLHIR RQKSVMPSER FVEPSRVAVA AKVHGGLDGA
51  AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*

```

m043/a043 100.0% identity in 129 aa overlap

```

              10      20      30      40      50      60
m043.pep    MVVSNQNIYAAGPSALLHIRRQKSVMPSEFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
              |||||
a043         MVVSNQNIYAAGPSALLHIRRQKSVMPSEFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
              10      20      30      40      50      60

              70      80      90     100     110     120
m043.pep    QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
              |||||
a043         QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
              70      80      90     100     110     120

              130
m043.pep    VVAAEGEAQX
              |||||
a043         VVAAEGEAQX
              130

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 181>:

```

g044.seq
1  ATGCTGCCCC ACCAGAGCGT CGAGTTCTTG CCACAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GCGGGCGGTT TTCCAGCCGT TGCGCTTCCA ACCGTGTATC
101 CAGTTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTTCAGCCC TTCGATAACG GCGGTCAGCT
201 CCATGCGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTTCG
251 CGGCTGCCGT AGCGCATTAA

```

This corresponds to the amino acid sequence <SEQ ID 182; ORF 044.ng>:

```

g044.pep
1  MLPDQSVEFL PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
51  GAAAFERFQP FDNGGQLHAV VGGLRFAAEK FFFAAVAH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 183>:

```

m044.seq
1  ATGCCGTCCG ACTAGAGCGT CGAGTTCTTT CCAGAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GGAGGCGGTT TTCCAGCCGT TGCGCTTCCA ACCGTGTATC
101 CAGTTTTTCCA TGCCATTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTTCAGTCC TTCGATGACG GCAGTCAGTT
201 CCATGCGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTTCG
251 TGGCTACCGT AGCGCAyTAa

```

This corresponds to the amino acid sequence <SEQ ID 184; ORF 044>:

```

m044.pep
1  MPSPDXSVEFF PEVVVFDGLF GGGFPAVALP TVYPVFHAIF DVLRVGADDD
51  GAAAFERFQS FDDGSQFHAV VGGLRFAAEK FFFVATVAH*

```



The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 185>:

```
a044.seq
  1 GTGCCGTCGG ACCAGCGCGT CGAGTTCTTT CCACAAGTCG TCGTTTTTGA
 51 CGGGCTGTTT GCGGCGCGTT TTCCAGCCGT TCGCTTCCA ACCGTGTATC
101 CAGTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTCAGTCC TTCGATGACG GCGGTCAGTT
201 CCATACGGTT GTTGGTGGTT TGCCTTCGC CGCCGAAAAG TTCTTTTTCG
251 TGGCTGCCGT AGCGCATTA
```

This corresponds to the amino acid sequence <SEQ ID 186; ORF 044.a>:

```
a044.pep
  1 VPSPQRVEFF PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
 51 GAAAFERFQS FDDGGQFHTV VGGLRFAAEK FFFVAVAH*
```

m044/a044 91.0% identity over a 89 aa overlap

	10	20	30	40	50	60
m044.pep	MPSDXSVEFFPEVVVFDGLFGGGFPVALPTVYPVFHAI FDLRVGADDDGAAAFERFQS					
	:           :           :           :					
a044	VPSPQRVEFFPQVVVFDGLFGGGFPVALPTVYPVFHAVFDL RVGADDDGAAAFERFQS					
	10	20	30	40	50	60
	70	80	90			
m044.pep	FDDGSQFHAVVGGLRFAAEKFF FVATVAHX					
	:       :           :           :					
a044	FDDGGQFHTVVGGLRFAAEKFF FVAVAHX					
	70	80	90			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 044 shows 86.5% identity over a 89 aa overlap with a predicted ORF (ORF 044.ng) from *N. gonorrhoeae*:

m044/g044

	10	20	30	40	50	60
m044.pep	MPSDXSVEFFPEVVVFDGLFGGGFPVALPTVYPVFHAI FDLRVGADDDGAAAFERFQS					
	:           :           :           :					
g044	MLPDQSVEFLPQVVVFDGLFGGGFPVALPTVYPVFHAVFDL RVGADDDGAAAFERFQP					
	10	20	30	40	50	60
	70	80	90			
m044.pep	FDDGSQFHAVVGGLRFAAEKFF FVATVAHX					
	:     :           :           :					
g044	FDNGGQLHAVVGGLRFAAEKFF FFAVAHX					
	70	80	90			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 187>:

```
g046.seq
  1 ATGTCGGCAA TGCTGCGTCC GACAAGCAGC CCGCCGCgcc gCGCCTGTAT
 51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CCGTAACGAG CTGTTCTGGGG
151 CTGATGGTTT CGGTTATGCC gaATATGGAA AGGCTGCCGt TTTcGTTGTT
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TtcgctGGAA CGGACGCGCG
251 CGATCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGagaGCGCG AGcagcaagt cggcatcttC
351 CgcgcggcgC Cgttataatg tgAAGGGGGA TGCGccgttg ccgaAAACGG
401 TTTGGacatc gaggcggctg CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451 TCGATAAcg9 TTACGTCGTT GTTGGTGATG GCGGCAAGGT TTTGCGCGAC
```

501 GGTAGAACCT ACCTGCCCCG TGCCTAAAAT GAGGATTTTC ACGGTATGGG  
551 TCGCCGGGTG A

This corresponds to the amino acid sequence <SEQ ID 188; ORF 046.ng>:

g046.pep

1 MSAMLRPTSS PRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG  
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT  
101 MLVSSLRESA SSKSASSAPA RYNVKGDAPL PKTVWTSRRL PVSCNAFSSM  
151 SITVTSLLMV AARFCATVEP TCPLPKMRIF TVWVAG\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 189>:

m046.seq

1 ATGTCGGCAA TGCTGCGTCC GACAAGCAST CCGC.r.sGC gCGcCTGTAT  
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC  
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCCGGG  
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT  
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG  
251 CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG  
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC  
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG  
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG  
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC  
501 GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG  
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 190; ORF 046>:

m046.pep

1 MSAMLRPTSX PXXRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG  
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT  
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM  
151 SITVTSLLMV AARFCATVEP TCPLPKMRIF TVWVAE\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 191>:

a046.seq

1 ATGTCGGCAA TGCTGCGTCC GACAAGCAGT CCGCCGCGCC GCGCCTGTAT  
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC  
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCCGGG  
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT  
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG  
251 CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG  
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC  
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG  
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG  
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC  
501 GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG  
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 192; ORF 046.a>:

a046.pep

1 MSAMLRPTSS PRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG  
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT  
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM  
151 SITVTSLLMV AARFCATVEP TCPLPKMRIF TVWVAE\*

m046/a046 98.4% identity over a 186 aa overlap

	10	20	30	40	50	60
m046.pep	MSAMLRPTSPXXRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG LMVSVMPNME					
a046	MSAMLRPTSSPPRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG LMVSVMPNME					

245

	10	20	30	40	50	60
m046.pep	70	80	90	100	110	120
	RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
a046	RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
	70	80	90	100	110	120
m046.pep	130	140	150	160	170	180
	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
a046	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
	130	140	150	160	170	180
m046.pep	TVWVAEX					
a046	TVWVAEX					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 046 shows 97.3% identity over a 185 aa overlap with a predicted ORF (ORF 046.ng) from *N. gonorrhoeae*:

m046/g046

	10	20	30	40	50	60
m046.pep	MSAMLRPTSXPPXRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME					
g046	MSAMLRPTSSPPRRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME					
	10	20	30	40	50	60
m046.pep	70	80	90	100	110	120
	RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
g046	RLPFSLFSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
	70	80	90	100	110	120
m046.pep	130	140	150	160	170	180
	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
g046	RYNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLVMAARFCATVEPTCPLPKMRIF					
	130	140	150	160	170	180
m046.pep	TVWVAEX					
g046	TVWVAGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 193>:

g047.seq

```

1  ATGGTCATCA TACAGGCGcg gcGCGGCGGG CTGCTTGTCG GACGCAGCAT
51  TGCCGACATC GCCCAAGATT TGCCCAGACGG GGCCGACTGC CAAATCTGCG
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCGTCATC
151 ATCGAAGGCG ACGAAATCCT GTTTGCCGCC GCCGCCGAAA ACATCGGGGC
201 GGTCATACCc gaATTGCGCC CCAAAGAAAC CAGCACCCGC CGCATCATGA
251 TTGCCGGCGG CGGCAACATc tgctACCGCC TCGCCAAGCA GCTCGAACAC

```

```

301 GCATAcacg TCAAAATCAT CGAATGCCGG CCGCGCcggtg ccgaATGGAT
351 AGCCGAAAAC ctcgAcaaCA CCCTCGTCCT GCAAGGTTCTG Gcaaccgacg
401 aAaccctgct cgAcaacgaa tacatcgacg aaatcgaCGT ATTCTGCGCC
451 CTGACCAACG ACGACGAAAG CAACATTatg tCCGCCCTTT TGGCGAAAAA
501 CCTcggcgCG AAGCgcgtca tcggCATCGT CAACCGCTCA AGCTACGTCG
551 ATTTGCTCGA AGGCAACAAA ATCGACATCG TCGTCTCCCC CCACCTCATC
601 ACCATCGGCT CGATACTCGC CCACATCCGG CGCGCGGACA TCGTTGCCGT
651 CCACCCCATC CGGCGCGGCA CGGCGGAAGC CATCGAAGTC GTCGCGCACG
701 GCGACAAAAA AACTTCCGCC ATCATCGGCA GGCGCATCAG CGGCATCAAA
751 TGGCCCGAAG GCTGCCACAT TGCCGCCGTC GTCCGCGCCG GAACCGGCGA
801 AACCATTATG GGACACCATA CCGAAACCGT CATCCAAGAC GGTGACCACA
851 TCATCTTTTT CGTCTCGCGC CGGCGCATCC TGAACGAAC TGGAGAAACTC
901 ATCCAAGTCA AAATGGGCTT TTTCCGATAA

```

This corresponds to the amino acid sequence <SEQ ID 194; ORF 047.ng>:

g047.pep

```

1  MVIIQARRGG LLVGRSIADI AQLDPDGADC QICAVYRNNR LIVPAPQTVI
51  IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI CYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDES NIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKK TSA IIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IQVKMGFFG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 195>:

m047.seq

```

1  ATGGTCATCA TACAGgCGcG C..syGCGGA sTGCTTGTCG GACGCAGCAT
51  TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCGTCATC
151 ATCGAAGGCG ACGAAATCCT ATTTGCCGCC GCCGCCGAAA ACATCGGCGC
201 GGTCATACCC GAATTGCGCC CCAAAGAAAC CCAAAGAAAC CAGCcCmgmm
251 GcATCATGAT TkCCGGCGGC GGCAACATCG GCTACCGTCT CGCCAAGCAG
301 CTCGAACACG CATAACAGT yAAAATCATC GAATGCCGGC CGCGCGTGC
351 CGAATGGATA GCCGAAAACC TCGACAACAC CCTCGTCyTG CAAGGTTCTGG
401 CAACCGACGA AACCCTGCTC GACAACGAAT ACATCGACGA AATCGACGTA
451 TTCTGCGCCC TGACCAACGA CGACGAAAGC AACATTATGT CCGCCCTTTT
501 GCGGAaAAAC CTCGGCGCGA AGCGCGTCAT CGGCATCGTC AACCGCTCAA
551 GCTACGTCGA TTTGCTCGAA GGCAACAAAA TCGACATCGT CGTCTCCCCC
601 CACCTCATCA CCATCGGCTC GATACTCGCC CACATCCGGC GCGGCGACAT
651 CGTTGCCGTC CACCCATCC GCGCGGCGAC GGCGGAAGCC ATCGAAGTCG
701 TCGCACACGG CGACAAAAA ACTTCCGCCA TCATCGGCAG GCGCATCAGC
751 GGCATCAAAT GGCCCGAAGG CTGCCACATT GCCGCCGTCG TCCGCGCCGG
801 AACC GGCGAA ACCATTATGG GACACCATA CGAAACCGTC ATCCAAGACG
851 GCGACCACAT CATCTTTTTC GTCTCGCGCC GCGCATCCT GAACGAAC TG
901 GAAAAACTCA TCCAGGTCAA AATGGGCTTT TTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 196; ORF 047>:

m047.pep

```

1  MVIIQARXXG XLVGRSIADI AQLDPDGADC QICAVYRNNR LIVPAPQTVI
51  IEGDEILFAA AAENIGAVIP ELRPKETORN QPXXIMIXGG GNIGYRLAKQ
101 LEHAYNVKII ECRPRRAEWI AENLDNTLV LQGSATDETL DNEYIDEIDV
151 FCALTNDDES NIMSALLAKN LGAKRVIGIV NRSSYVDLLE GNKIDIVVSP
201 HLITIGSILA HIRRGDIVAV HPIRRGTAE AIEVVAHGD K TSAIIGRRIS
251 GIKWPEGCHI AAVVRAGTGE TIMGHHTETV IQDGDHIIFF VSRRRILNEL
301 EKLIQVKMGF FG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 197>:

a047.seq

```

1  ATGGTCATCA TACAGGCGCG GCGCGGCGGA CTGCTTGTCG GACGCAGCAT

```

```
51  TGCCGACATC  GCCCAAGATT  TGCCCGACGG  GGCCGACTGC  CAAATCTGCG
101 CCGTTTACCG  CAACAACCGC  CTCATCGTCC  CCGCGCCGCA  AACCGTCATC
151 ATCGAAGGCG  ACGAAATCCT  ATTTGCCGCC  GCCGCCGAAA  ACATCGGCGC
201 GGTATACCC  GAATTGCGCC  CCAAAGAAAC  CAGCACCCGC  CGCATCATGA
251 TTGCCGGCGG  CGGCAACATC  GGCTACCGTC  TCGCCAAGCA  GCTCGAACAC
301 GCATACAACG  TCAAATCAT  CGAATGCCGG  CCGCGCCGTG  CCGAATGGAT
351 AGCCGAAAAC  CTCGACAACA  CCCTCGTCCT  GCAAGGTTTC  GCAACCGACG
401 AAACCTGCT  CGACAACGAA  TACATCGACG  AAATCGACGT  ATTCTGCGCC
451 CTGACCAACG  ACGACGAAAG  CAACATTATG  TCCGCCCTTT  TGGCGAAAAA
501 CCTCGGCGCG  AAGCGCGTCA  TCGGCATCGT  CAACCGCTCA  AGCTACGTCG
551 ATTTGCTCGA  AGGCAACAAA  ATCGACATCG  TCGTCTCCCC  CCACCTCATC
601 ACCATCGGCT  CGATACTCGC  CCACATCCGG  CGCGGCGACA  TCGTTGCCGT
651 CCACCCCATC  CGGCGCGGCA  CGGCGGAAGC  CATCGAAGTC  GTCGCACACG
701 GCGCAAAAA  AACTTCCGCC  ATCATCGGCA  GGCGCATCAG  CGGCATCAAA
751 TGGCCGAAG  GCTGCCACAT  TGCCGCCGTC  GTCCGCGCCG  GAACCGGCGA
801 AACCATTATG  GGACACCATA  CCGAAACCGT  CATCCAAGAC  GGCGACCACA
851 TCATCTTTT  CGTCTCGCGC  CGGCGCATCC  TGAACGAACT  GGAAAACTC
901 ATCCAAGTCA  AAATGGGCTT  TTTCGATAA
```

This corresponds to the amino acid sequence <SEQ ID 198; ORF 047.a>:

```
a047.pep
1  MVIIQARRGG LLVGRSIADI AQDLDPGADC QICAVYRNNR LIVPAPQTVI
51  IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI GYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDSNIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTS IIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IQVKMGFFG*
```

m047/a047 96.5% identity over a 312 aa overlap

```
10      20      30      40      50      60
m047.pep  MVIIQARXXGXLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047      MVIIQARRGGLLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA
          10      20      30      40      50      60

70      80      90      100     110     120
m047.pep  AAENIGAVIPELRPKETQRNQPXXIMIXGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047      AAENIGAVIPELRPKETSTR---IMIAGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI
          70      80      90      100     110

130     140     150     160     170     180
m047.pep  AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047      AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV
          120     130     140     150     160     170

190     200     210     220     230     240
m047.pep  NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047      NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK
          180     190     200     210     220     230

250     260     270     280     290     300
m047.pep  TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a047      TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
          240     250     260     270     280     290

310
m047.pep  EKLIQVKMGFFGX
          | | | | | | | |
```

a047            EKLIQVKMGFFGX  
                 300            310

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 047 shows 96.2% identity over a 312 aa overlap with a predicted ORF (ORF 047.ng) from *N. gonorrhoeae*:

m047/g045

m047.pep	MVIIQARXXGXLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA	60
g047	MVIIQARRGGLLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA	60
m047.pep	AAENIGAVIPELRPKETQRNQPXXIMIXGGNIGYRLAKQLEHAYNVKIIIECRPRAEWI	120
g047	AAENIGAVIPELRPKETSTR---IMIAGGGNICYRLAKQLEHAYNVKIIIECRPRAEWI	117
m047.pep	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	180
g047	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	177
m047.pep	NRSSVVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	240
g047	NRSSVVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	237
m047.pep	TSIIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRLNEL	300
g047	TSIIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRLNEL	297
m047.pep	EKLIQVKMGFFGX    313	
g047	EKLIQVKMGFFGX    310	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 199>:

g048.seq

1	ATGCTCGACA	AAGGCGAGGA	GTTGCCCGTC	GATTTACCA	ACCGCCTGAT
51	TTACTACGTC	ggcCCcgTCG	ATCCGGTCGG	CGATGAAGTC	GTCGGTCCCG
101	CAGGTCCGAC	CACAGCCACC	CGCATGGACA	AATTTACCCG	CCAAATGCTC
151	AAACAAACCG	GCCTCTTGGG	CATGATCGGC	AAATCCGagc	gcgGcgcggc
201	cacctGCGAA	GccatCGCCG	ACAACAAGGC	CGTGTACCTC	ATGGCAGTCG
251	GCGGCGCGGC	ATACCTCGTG	GCAAAAGCCA	TCAAATCTTC	CAAAGTCTTG
301	GCGTTCCCCG	AATTGGGTAT	GGAAGCCGTT	TACGAATTTG	AAGTCAAAGA
351	TATGCCCCGA	ACCGTCGCCG	TGGACAGCAA	AGGCGAATCC	ATCCACGCCA
401	CCGCCCCGCG	CAAATGGCAG	GCGAAAATCG	GCATCATCCC	CGTCGAGTCT
451	TGA				

This corresponds to the amino acid sequence <SEQ ID 200; ORF 048.ng>:

g048.pep

1	MLDKGEELPV	DFTNRLIYYV	GPVDPVGDEV	VGPAGPTTAT	RMDKFTRQML
51	KQTGLLGMI	GKSERGAATCE	AIADNKAVYL	MAVGGAAYLV	AKAIKSSKVL
101	AFPELGMEAV	YEFVVKDMPV	TVAVDSKGES	IHATAPRKWQ	AKIGIIPVES
151	*				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 201>:

m048.seq

1	ATGCTCAACA	AAGGCGAAGA	ATTGCCCGTC	GATTTACCA	ACCGCCTGAT
51	TTACTACGTC	GGCCCCGTCG	ATCCGGTCGG	CGATGAAGTC	GTCGGTCCCG

```

101 CAGGTCCGAC CACAGCCACC CGCATGGACA AATTCACCCG CCAAATGCTC
151 GAACAAACCG ACCTCTTGGG CATGATCGGC AAATCCGAGC GCGGCGTGCC
201 CACCTGCGAA GCCATCGCCG ACAACAAAGC CGTGTACCTC ATGGCAGTCG
251 GCGGCGCGGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
301 GCGTTCCCCG AATTGGGCAT GGAAGCCATT TACGAATTG AAGTCAAAGA
351 CATGCCCGTA ACCGTCGCCG TAGATAGCAA AGGCGAATCC ATCCACGCCA
401 CCGCCCCGCG CAAATGGCAG GCGAAAATCG GCATCATCCC CGTCGAATCT
451 TGA

```

This corresponds to the amino acid sequence <SEQ ID 202; ORF 048>:

```

m048.pep
  1 MLNKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML
 51 EQTDLLGMIG KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101 AFPPELGMEAI YEFVVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
151 *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 203>:

```

a048.seq
  1 ATGCTCGACA AAGGCGAAGA ATTGCCCGTC GATTTACCA ACCGCCTGAT
 51 TTACTACGTC GGCCCCGTCG ATCCGGTCGG CGACGAAATC GTCGGCCCAG
101 CAGGTCCGAC CACCGCCACC CGCATGGACA AATTCACCCG CCAAATGCTC
151 GAACAAACCG ACCTCTTGGG CATGATCGGC AAATCCGAGC GCGGCGCGGC
201 CACCTGCGAA GCCATCGCCG ACAACAAAGC CGTGTACCTC ATGGCAGTCG
251 GCGGCGCGGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
301 GCGTTCCCCG AATTGGGCAT GGAAGCCATT TACGAATTG AAGTCAAAGA
351 CATGCCCGTA ACCGTCGCCG TAGACAGCAA AGGCGAATCC ATCCACGCCA
401 CCGCCCCGCC CCAATGGCAG GCGAAAATCG GCATCATCCC CGTCAAATCT
451 TGA

```

This corresponds to the amino acid sequence <SEQ ID 204; ORF 048.a>:

```

a048.pep
  1 MLDKGEELPV DFTNRLIYYV GPVDPVGDEI VGPAGPTTAT RMDKFTRQML
 51 EQTDLLGMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101 AFPPELGMEAI YEFVVKDMPV TVAVDSKGES IHATAPPQWQ AKIGIIPVKS
151 *

```

**m048/a048** 96.0% identity over a 150 aa overlap

	10	20	30	40	50	60
m048.pep	MLNKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML EQTDLLGMIG					
a048	MLDKGEELPV DFTNRLIYYV GPVDPVGDEI VGPAGPTTAT RMDKFTRQML EQTDLLGMIG					
	10	20	30	40	50	60
m048.pep	KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL AFPPELGMEAI YEFVVKDMPV					
a048	KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL AFPPELGMEAI YEFVVKDMPV					
	70	80	90	100	110	120
m048.pep	KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL AFPPELGMEAI YEFVVKDMPV					
a048	KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL AFPPELGMEAI YEFVVKDMPV					
	70	80	90	100	110	120
m048.pep	TVAVDSKGES IHATAPRKWQ AKIGIIPVESX					
a048	TVAVDSKGES IHATAPPQWQ AKIGIIPVKSX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 048 shows 96.4% identity over a 150 aa overlap with a predicted ORF (ORF 048.ng) from *N. gonorrhoeae*:

m048/g048

```

      10      20      30      40      50      60
m048.pep MLNKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIG
||:|||||||||||||||||||||||||||||||||||||:|||||
g048      MLDKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLKQTGLLGMIG
      10      20      30      40      50      60

      70      80      90      100     110     120
m048.pep KSERGVATCEAIADNKAVYLMVAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPV
||||:|||||||||||||||||||||||||||||||||:|||||
g048      KSERGAATCEAIADNKAVYLMVAVGGAAYLVAKAIKSSKVLAFPELGMEAVYEFVVKDMPV
      70      80      90      100     110     120

      130     140     150
m048.pep TVAVDSKGESIHATAPRKWQAKIGIIPVESX
|||||||||||||||||
g048      TVAVDSKGESIHATAPRKWQAKIGIIPVESX
      130     140     150

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 205>:

g049.seq

```
1 ATGCGGGGCGC AGGCGTTTGA TCAACCGTTC GGTGAGCTCC TGTTCGGACA
51 GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG GATATTGATT
101 TGGACGGGCA TCAACGCCTC TTCCGCACCG CCTTCGCCGT TTTCCGCAAC
151 CCCGTCTGCC GCCGTACCGG ATTCTGCCGC ATCGGCGTTT TCCCCGCCCT
201 CAATCTGTGC GGTTTCAAAT TCGGCACGT CTTTTTTGGC ATCGAACCGG
251 ATTCTCCGCC GCGATTTCGAT GTGTTTTTCC GAAAccggca tTTGCAGGGA
301 AGCCTgcg9c TTGAGCCAGT TTTCTGAAG GACGATCATC GGGTCGGTTT
351 CGACTTCCTC GCCGCAATCG GCAACGGCgc tGTTGTGTTT TTCCTGCCAT
401 TTCTTCAGAT ACGCCTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 206; ORF 049.ng>:

g049.pep

```
1  MRAQAFDQPF  GQLLFGQAEH  FAPVDGFRVQ  DIDLDGHQRL  FRTAFAVFRN
51  PVCRRTGFCR  IGVFPALNLC  GFKFGTVFFG  IEPDSPPRFD  VFFRNRHLQG
101 SLRVEPVFLK  DDHRVGDFDL  AAIGNGAVVF  FLPLQLIRL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 207>:

m049.seq (partial)

1	ATGCGGGCGC	AGGCGTTTGA	TCAGCCGTTT	GGTCAGCTCC	TGTTCGGACA
51	GGCAGAACAC	TTCGCGCCGG	TTGACGGCTT	TCGGGTTT	CAGGATATTGATT
101	TGGACGGGCA	TCAACGTTTC	TTCCGCATCG	TTTTCCCCGT	TTTCCGAAAC
151	CGCCGGCTCA	TTCGTGCCGG	ATTCTGCCTC	GTCCGGCGTTT	TCCCCGCTTT
201	CAATCTGTCC	GGTTTCAAAT	TCGACACTGT	CTTTTTTGGT	ATCAAACCGG
251	ATTCTCCGCC	GCGATTTCGAT	GTGTTTTTCC	GAAACCGACA	TTTGCAGGGA
301	AGCCTGCGCG	TTGAGCCAGT	TTTCTGAAG	GACGATCATC	GGGTTCGGTTT
351	CGACTTCCTC	GCCGCAATCG	GCAACGGCGG	CATTGTGTTC	CTCCTGCCAT
401	TTTTTCAGAT	ACGCCTT...			

This corresponds to the amino acid sequence <SEQ ID 208; ORF 049>:

m049.pep (partial)

1	MRAQAFDQPF	QQLLFGQAEH	FAPVDGFRVQ	DIDLGDGHQRF	FRIVFPVFRN
51	RRLIRAGFCL	VGVFPAFNLS	GFKFDTVFFG	IKPDSPPRFD	VFFRNRHLQG
101	SLRVEPVFLK	DDHRVGDFDL	AAINGGGIVF	LLPFFQIRL...	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 209>:

a049.seq



```

1  ATGCGGGCGC AGGCGTTTGA TCAGCCGTTT GGTCTAGCTCC TGTTCGGACA
51  GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG AATATTGATT
101 TGGACGGGCA TCAACGCTTC TTCCGCACCG CCTTCGCCGT TTTCCGCAAC
151 CCCGTCTGCC GCCGTACCG ATTCTGCCG ATCGGCGTTT TCCCGCCTT
201 CAATCTGTCC GGTTCCTAAAT TCGGCACTGT CTTTTTTGGC ATCAAACCGG
251 ATTCTCCGCC GCGATTTCGAT GTGTTTTTCC GAAACCGACA TTTGCAGGGA
301 AGCCTGCGCG TTGAGCCAGT TTTCTGAAG GACGATCATC GGGTCGGTTT
351 CGACTTCCTC GCCGCAATCG GCAACGGCGG CATTGTGTTC CTCCTGCCAT
401 TTTTTCAGAT ACGCCTT

```

This corresponds to the amino acid sequence <SEQ ID 210; ORF 049.a>:

```

a049.pep
1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ NIDLDGHQRF FRTAFAVERN
51  PVCRRTRFCR IGVFPAFNLS GFKFGTVFFG IKPDSPPRFD VFFNRHLQG
101 SLRVEPVFLK DDHRVGFDL AAIGNGGIVF LLPFFQIRL

```

m049/a049 90.6% identity over a 139 aa overlap

```

              10      20      30      40      50      60
m049.pep      MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLDGHQRFRRIVFPVFRNRRLIRAGFCL
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a049           MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQNIDLDGHQRFRTAFAVERNPVCRRTRFCR
              10      20      30      40      50      60

              70      80      90      100     110     120
m049.pep      VGVFPAFNLSGFKFDTVFFGIKPDSPPRFDVFFNRHLQGSRLRVEPVFLKDDHRVGFDL
              :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a049           IGVFPAFNLSGFKFGTVFFGIKPDSPPRFDVFFNRHLQGSRLRVEPVFLKDDHRVGFDL
              70      80      90      100     110     120

              130     139
m049.pep      AAIGNGGIVFLLPFFQIRL
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a049           AAIGNGGIVFLLPFFQIRL
              130

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 049 shows 86.3% identity over a 139 aa overlap with a predicted ORF (ORF 049.ng) from *N. gonorrhoeae*:

```

m049/g049
              10      20      30      40      50      60
m049.pep      MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLDGHQRFRRIVFPVFRNRRLIRAGFCL
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g049           MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLDGHQRLFRRTAFAVERNPVCRRTGFCR
              10      20      30      40      50      60

              70      80      90      100     110     120
m049.pep      VGVFPAFNLSGFKFDTVFFGIKPDSPPRFDVFFNRHLQGSRLRVEPVFLKDDHRVGFDL
              :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g049           IGVFPALNLCGFKFGTVFFGIEPDSPPRFDVFFNRHLQGSRLRVEPVFLKDDHRVGFDL
              70      80      90      100     110     120

              130     139
m049.pep      AAIGNGGIVFLLPFFQIRL
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g049           AAIGNGAVVFFLPFLQIRLX
              130     140

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 211>:

g050.seq

```

1 atgggCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGg
51 cacgcccGAA AAAGccgtgt TGATGGcaaa AGAATCCCTG ATGAGCCACA
101 TCGAcAtCca aGaATTGCAG GAAAAAGCCG CGTccg999c ggaattgtcc
151 accaccgaAG ccCTGCGCCT cGAACTCTTT GAAAAAGTCA ACGCGCTGGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAT GCCGCCTCCA AACCGATTGC CATGATTCCC
301 AACTGTGCcg ccacCCGcca cgtcgAATTT GAATTGgACG GCTCAGgtcc
351 TGTCGAactc acgccGCcgc gtgtCGAAGA CTGA

```

This corresponds to the amino acid sequence <SEQ ID 212; ORF 050.ng>:

g050.pep

```

1 MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51 TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVED*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 213>:

m050.seq

```

1 ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGTATCG GCATCGGCGG
51 C..agCCgAA AAAGCCGTGC TGATGGCAAA AGAGTCCCTG ATGAGCCACA
101 TCGACATTCA AGAATTGCAG GAAAAGGCCG CGTCCGGCGC GgAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTGGG
201 CATCGGCGCA CAAGGCTTGG GCGGACTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA TCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCC
351 TGTCGAATC ACGCCGCCGC GCGTCGAAGA TGGCCCGATT TGA

```

This corresponds to the amino acid sequence <SEQ ID 214; ORF 050>:

m050.pep

```

1 MGAGWCPPGI LGIGIGGXAE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51 TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDGPI *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 215>:

a050.seq

```

1 ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGG
51 TACGCCCCGAA AAAGCCGTGT TGATGGCGAA AGAATCCCTG ATGAGCCACA
101 TCGACATCCA AGAATTGCAG GAAAAGCCG CGTCCGGCGC GGAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTAGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCC
351 TGTCGAATC ACGCCGCCGC GCGTCGAAGA CTGGCCC

```

This corresponds to the amino acid sequence <SEQ ID 216; ORF 050.a>:

a050.pep

```

1 MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51 TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDWP

```

m050/a050 97.7% identity over a 129 aa overlap

```

          10      20      30      40      50      60
m050.pep  MGAGWCPPGILGIGIGGXAEKAVLMAKESLSMSHIDIQELQ EKAASGAELSTTEALRLELF
          |||||:|||||
a050      MGAGWCPPGILGIGIGGTPEKAVLMAKESLSMSHIDIQELQ EKAASGAELSTTEALRLELF

```

	10	20	30	40	50	60
	70	80	90	100	110	120
m050.pep	EKVNALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
a050	EKVNALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
a050	TPPRVEDWP					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 050 shows 98.4% identity over a 127 aa overlap with a predicted ORF (ORF 050.ng) from *N. gonorrhoeae*:

m050/g050

	10	20	30	40	50	60
m050.pep	MGAGWCPPGILGIGIGGXAEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLLELF					
g050	MGAGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLLELF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m050.pep	EKVNALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
g050	EKVNALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
g050	TPPRVEDX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 217>:

g050-1.seq

```

1  ATGACCGTTA TCAAGCAAGA AGACTTTATT CAAAGTATCT GCGATGCCTT
51  CCAATTCATC AGCTACTACC ATCCAAAAGA CTACATCGAC GCGCTTTATA
101 AGGCGTGGCA GAAGGAAGAA AATCCCGCCG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGTAT GTGTGCCGAA AACAACCGCC CCATCTGCCA
201 AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG GATGTGCAAT
251 GGGATGCGGA CATGAGCGTG GAAAAGATGG TTAACGAAGG CGTACGCCGC
301 GCCTACACTT GGGGAAGGCAA CACCCTGCGC GCTTCCGTCC TCGCCGATCC
351 GGCCGGCAAA CGCCAAAACA CCAAAGACAA CACCCCGGCC GTCATCCACA
401 TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GGCGGCTCTG AAAACAAATC CAAACTCGCT ATGCTCAACC CTTCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACGATGGGC GCGGGCTGGT
551 GTCCCTCCCG CATCTTGGGC ATCGGCATCG GCGGCACgC CGAAAAAGCC
601 GTGTTGATGG cgaAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
651 GCAGGAAAAA GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTAAAAAG GTCAACGCGC TGGGCATCGG CGCGCAAGGC
751 TTGGGCGGTC TGACCAACCGT GTTGGACGTG AAAATCCTCG ATTACCCGAC
801 CCATGCCGCC TCCAAACCGA TTGCCATGAT TCCCAACTGT GCCGCCACCC
851 GCCACGTCGA ATTTGAATTG GACGGCTCAG GTCTGTGCA ACTCACGCCG
901 CCGCGCGTCG AAGACTGACC CGATCTGACT TACAGCCCGG ACAACGGCAA
951 ACGCGTCGAT GTCGATAAGC TGACCAAAGA AGAAGTGGCA AGCTGGAAAA
1001 CGGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACC GGCGCATGCC
1051 CCGCACAAAC GCCTCGTCAA TATGCTCGAC AAAGGCGAGG AGTTGCCCGT
1101 CGATTTCACC AACCGCCTGA TTTACTACGT CGGCCCGGTC GATCCGGTCG
1151 GCGATGAAGT CGTCGGTCCC GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201 AAATTTACCC GCCAAATGCT CAAACAAACC GGCTCTTGG GCATGATCGG

```

```

1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAGG
1301 CCGTGTACCT CATGGCAGTC GGCGGCGCGG CATACCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GCGGTTCCCC GAATTGGGTA TGGAAAGCCGT
1401 TTACGAATTT GAAGTCAAAG ATATGCCCGT AACCGTCGCC GTGGACAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCGC GCAAATGGCA GGCAGAAATC
1501 GGCATCATCC CCGTCGAGTC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 218; ORF 050-1.ng>:

g050-1.pep

```

1  MTVIKQEDFI QSIDAFQFI SYHHPKDYID ALYKAWQKEE NPAAKDAMTQ
51  ILVNSRMCAE NNRPICQDTG IATVFLKVGM DVQWDADMSV EKMVNEGVRR
101 AYTWEQNTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGG KVEVTCAAKG
151 GGSSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELF EK VNALGIGAQG
251 LGGLTTVLDV KILDYPTTHA SKPIAMIPNC AATRHVEFEL DSGSPVELTP
301 PRVED*PDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLL NGKILTGRDA
351 AHKRLVNMLD KGEELPVDFE NRIIYVGPV DPVGDEVVGP AGPTTATRMD
401 KFTTRQMLKQT GLLGMIGKSE RGAATCEAIA DNKAVYLMV GGAAYLVAKA
451 IKSSKVLAFP ELGMEAVYEF EVKDMPVTV VDSKGESIHA TAPRKWQAKI
501 GIIPVES*

```

g050-1/p14407

```

sp|P14407|FUMB_ECOLI FUMARATE HYDRATASE CLASS I, ANAEROBIC (FUMARASE)
>gi|280063|pir|B44511 fumarate hydratase (EC 4.2.1.2) fumB, iron-dependent - Escherichia coli
>gi|146048 (M27058) anaerobic class I fumarase (EC 4.2.1.2) [Escherichia coli] Length = 548
Score = 172 bits (432), Expect = 4e-42
Identities = 138/488 (28%), Positives = 216/488 (43%), Gaps = 22/488 (4%)

```

```

Query: 11  QSIDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAENNRPICQDTG 70
          Q+  DA  +  H K  L+  E  +  K  Q L NS + A+  P CQDTG
Sbjct: 53  QAFHDASFMLRPAHQKQVAAILHDPEASEND---KYVALQFLRNSEIAAKGVLPCTCQDTG 109

Query: 71  IATVFLKVGMVDVQWDADMSVEKMVNEGVRRAYTWEGNTLRASVLADPAGKRQNTKDNTPA 130
          A +  K G V W  E+ +++GV  Y E N  +  A  K NT N PA
Sbjct: 110 TAIIVGKKQQRV-WTGGGD-EETLSKGVYNTYI-EDNLRYSQNAALDMYKEVNTGTNLPA 166

Query: 131 VIHMSIVPGGKVEVTCAAKGGGSENKSKL-----AMLNPSDNIVDWVLKTIPTMGAGWCP 185
          I +  V G + +  C AKGGGS NK+ L  A+L P  + +++++ + T+G  CP
Sbjct: 167 QIDLYAVDGDDEYKFLCVAKGGGSANKTYLYQETKALLTPG-KLKNFLVEKMRTLGTAAACP 225

Query: 186 PXXXXXXXXTPEKAVLMAKESLSHIDIQELQEKAAASGAELSTTEALRLELF EKVNXXX 245
          P  T + L +  +H  EL +  +  L EL E+
Sbjct: 226 PYHIAFVIGGTS AETNLKTVKLASAHY-YDELPTEGNEHQAFRDVQLEQELLEEAQKLG 284

Query: 246 XXXXXXXXXXTTVLDVKILDYPTTHAASKPIAMIPNCAATRHVEFELDGGG----PVELTP 301
          D++++ P H AS P+ M +C+A R+++ +++ G  +E P
Sbjct: 285 LGAQFGGKYFAH-DIRVIRLPRHGASCFVGMGVSCSADRNIKAKINREGIWKLEHNPG 343

Query: 302 RVEDXPDLTYSPDNGKRVDVDKLTKE---EVASWKTGDVLLNGKILTGRDAAHKRLVNM 358
          +  +VD+++ KE  +++ +  L L G I+ GRD AH +L +
Sbjct: 344 QYIPQELRQAGEGEAVKVDLNRPMKEILAQLSQYFVSTRSLTGTIIVGRDIAHAKLKL 403

Query: 359 LDKGEELPVDFEFTNRIIYXXXXXXXXXXXXXXXXXTATRMDKFTTRQMLKQTGLLGMIGK 418
          +D G+ELP  + IYY  TTA RMD +  +  G + M+ K
Sbjct: 404 IDAGKELPQYIKDHPYIYAGPAKTPAGYPSGSLGPTTAGRMDSYVDLLQSHGGSMIMLAK 463

Query: 419 SERGAATCEAIAADNKAVYLMVAVGG-AAYLVAKAIKSSKVLAFPPELGMEAVYEFVKDMPV 477
          R  +A  +  YL ++GG AA L ++IK  + +A+PELGMEA+++ EV+D P
Sbjct: 464 GNRSQQVTDACHKHGGFYLGSI GGPAVLAQOSIKHLECVAYPELGMEAIIWKIEVEDFPA 523

Query: 478 TVAVDSKG 485
          + VD KG
Sbjct: 524 FILVDDKG 531

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 219>:

m050-1.seq

```

1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
51  CCAATTCATC AGCTACTATC ATCCCAAAGA CTACATCGAC GCGCTTTATA
101 AGGCGTGGCA GAAGGAAGAA AATCCTGCGG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGTAT GTGTGCGGAA AACACCGGCC CCATCTGCCA
201 AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG AACGTCCAAT
251 GGGATGCGGA CATGAGCGTG GAAGAGATGG TTAACGAAGG CGTACGCCGC

```

```
301 GCCTACACTT GGAAGGCAA TACGCTGCGC GCTTCCGTCC TCGCCGATCC
351 GGCCGGCAAAA CGCCAAAACA CCAAAGACAA CACCCCGGCC GTCATCCATA
401 TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GCGGCTCTG AAAACAAATC CAAACTCGCC ATGCTCAATC CTTCCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACCATGGGC GCGGGCTGGT
551 GTCTCTCCCG CATCTTGGGT ATCGGCATCG GCGGCACGCC CGAAAAAGCC
601 GTGCTGATGG CAAAAGAGTC CCTGATGAGC CACATCGACA TTCAAGAATT
651 GCAGGAAAAG GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAAAA GTCAACGCGC TGGGCATCGG CGCACAAAGG
751 TTGGGCGGAC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTATCCGAC
801 CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAAGTCG GCCGCCACCC
851 GCCACGTCGA ATTTGAATTG GACGGCTCAG GCCCTGTCGA ACTCAGCCCG
901 CCGCGCGTCG AAGACTGGCC CGATTGACT TACAGCCCCG ACAACGGCAA
951 ACGGCTCGAT GTCGACAAGC TGACCAAGA AGAAGTGCCA AGCTGAAAAA
1001 CCGCGACGAT ATTGCTGTTG AACGGCAAAA TCCTCACCAG CCGCGATGCC
1051 GCACACAAAC GCCTCGTCGA TATGCTCAAC AAAGGCGAAG AATTGCCCGT
1101 CGATTTCACC AACCGCTCGA TTTACTACGT CGGCCCGCTC GATCCGGTCG
1151 CCGATGAAGT CGTCGGTCCG GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201 AAATTCACCC GCCAAATGCT CGAACAACC GACCTCTTGG GCATGATCGG
1251 CAAATCCGAG CGCGGCGTGG CCACCTGCGA AGCCATCGCC GACAACAAAG
1301 CCGGTACCTT CATGGCAGTC GCGGCGCGCG CGTATCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GCGGTTCCCG GAATTGGGCA TGGAAGCCAT
1401 TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGATAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCGCG GCAAATGGCA GGCGAAAATC
1501 GGCATCATCC CCGTCGAATC TTGA
```

This corresponds to the amino acid sequence <SEQ ID 220; ORF 050-1>:

m050-1.pep

```
1  MTVIKQEDFI QSIDAFQFI SYHPKDYID ALYKAWQKEE NPAAKDAMTQ
51  ILVNSRMCAE NNRPICQDTG IATVFLKVG M NVQWDADMSV EEMVNEGVR
101 AYTWEGNTLR ASVLADPAGK RONTKDNTPA VIHMSIVPGG KVEVTCAAKG
151 GGSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELF EK VNALGIGA QG
251 LGGTLTVLDV KILDYPTHA A SKPIAMIPNC AATRHVEF EL DSGSPVELTP
301 PRVEDWPDLT YSPDNKGRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
351 AHKRLVMDLN KGEELPVDF T NRLIYVGPV DPVGDEVVGP AGPTTATRMD
401 KFTRQMLEQT DLLGMIGKSE RGVATCEAIA DNKAVYLM AV GGAAYLVAKA
451 IKSSKVLAFP ELGMEAIYEF EVKDMPTVTA VDSKGESIHA TAPRKWQAKI
501 GIIPVES*
```

m050-1/g050-1 98.2% identity in 507 aa overlap

```
10      20      30      40      50      60
m050-1.pep MTVIKQEDFIQSIDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
|||||
g050-1      MTVIKQEDFIQSIDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
10      20      30      40      50      60

70      80      90      100     110     120
m050-1.pep NNRPICQDTGIATVFLKVMNVQWDADMSVEEMVNEGVRRAITWEGNLTIRASVLADPAGK
|||||
g050-1      NNRPICQDTGIATVFLKVMNVQWDADMSVEEMVNEGVRRAITWEGNLTIRASVLADPAGK
70      80      90      100     110     120

130     140     150     160     170     180
m050-1.pep RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGSENKSKLMLNPSDNIVDWVLKTIPTMG
|||||
g050-1      RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGSENKSKLMLNPSDNIVDWVLKTIPTMG
130     140     150     160     170     180

190     200     210     220     230     240
m050-1.pep AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELQEKAAASGAELSTTEALRLELF EK
|||||
g050-1      AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELQEKAAASGAELSTTEALRLELF EK
190     200     210     220     230     240

250     260     270     280     290     300
m050-1.pep VNALGIGA QGLGGLTTVLVDVKILDYPTHAASKPIAMIPNCAATRHVEFELD GSGPVELTP
|||||
g050-1      VNALGIGA QGLGGLTTVLVDVKILDYPTHAASKPIAMIPNCAATRHVEFELD GSGPVELTP
250     260     270     280     290     300
```

```

          310      320      330      340      350      360
m050-1.pep PRVEDWPDLTYS PDNGKRV DDKLTKEEVASWKTGDVLLNGKILTRDAAHKRLV DMLN
          |||||
g050-1      PRVEDXPDLTYS PDNGKRV DDKLTKEEVASWKTGDVLLNGKILTRDAAHKRLV DMLN
          310      320      330      340      350      360

          370      380      390      400      410      420
m050-1.pep KGEELPVDF TNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDL LGMIGKSE
          |||||
g050-1      KGEELPVDF TNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLKQTGL LGMIGKSE
          370      380      390      400      410      420

          430      440      450      460      470      480
m050-1.pep RGVATCEA IADNKAVY LMAVGG AAYLVAKAIKSSKVLAFPELGMEAIYEFV KDMPTVTA
          ||:|||||
g050-1      RGAATCEA IADNKAVY LMAVGG AAYLVAKAIKSSKVLAFPELGMEAVYEFV KDMPTVTA
          430      440      450      460      470      480

          490      500
m050-1.pep VDSKGESI HATAPRKWQAKIGIIPVESX
          |||||
g050-1      VDSKGESI HATAPRKWQAKIGIIPVESX
          490      500

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 221>:

a050-1.seq

```

1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATT TCGCATGCCTT
51  CCAATTCATC AGCTACTACC ATCCCAAAGA CTACATCGAC GCGCTTTATA
101 AGGCGTGGCA GAAGGAAGAA AACCCCGCCG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGCAT GTGTGCCGAA AACAAACGCC CCATCTGCCA
201 AGATACCGGT ATCGCGACCG TGTTTTTGAA AGTCGGTATG GATGTGCAAT
251 GGGATGCAGA CATGAGCGTC GAAGAGATGG TTAACGAAGG CGTGCGCCGC
301 GCCTACACTT GGAAGGCAA TACGCTGCGC GCTTCCGTTT TCGCCGACCC
351 CGCCGGCAAA CGCCAAAATA CCAAAGACAA CACGCCGCCG GTCATCCATA
401 TGAGCATCGT GCCGGGCGAC AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GGCGGTTCTG AAAACAAATC CAAACTCGCC ATGCTCAACC CTTCGACAA
501 CATCGTCGAT TGGGTATTGA AAACATTCC GACCATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGC ATCGGCATCG GCGGTACGCC CGAAAAAGCC
601 GTGTTGATGG CGAAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
651 GCAGGAAAAA GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAAAA GTCAACGCGC TAGGCATCGG CGCGCAAGGC
751 TTGGGCGGTC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTACCCGAC
801 CCACGCCGCC TCCAACCGA TTGCCATGAT TCCGAACGTC GCCGCCACCC
851 GCCACGTCGA ATTTGAATTG GACGCTCAG GCCCTGTCGA ACTCACGCCG
901 CCGCGCGTCG AAGACTGGCC CGATTTGACT TACAGCCCCG ACAACGGCAA
951 ACGCGTCGAT GTCGACAAGC TGACCAAAGA AGAAGTGGCA AGCTGAAAA
1001 CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACC GGCGGATGCC
1051 GCACACAAAC GCCTCGTCGA TATGCTCGAC AAAGGCGAAG AATTGCCCGT
1101 CGATTTCACC AACCGCCTGA TTTACTACGT CGGCCCGGTC GATCCGGTGC
1151 GCGACGAAAT CGTCGGCCCA GCAGGTCCGA CCACCGCCAC CCGCATGGAC
1201 AAATTCACCC GCCAAATGCT CGAACAAACC GACCTCTTGG GCATGATCGG
1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAAG
1301 CCGTGTACCT CATGGCAGTC GCGGCGCGCG CGTATCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GCGGTTCCCC GAATTGGGCA TGGAAAGCCAT
1401 TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGACAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCGCG CCAATGGCA GGCGAAAAATC
1501 GGCATCATCC CCGTCAAATC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 222; ORF 050-1.a>:

a050-1.pep

```

1  MTVIKQEDFI QSIDAFQFI SYHHPKDYID ALYKAWQKEE NPAAKDAMTQ
51  ILVNSRMCAE NNRPICQDTG IATVFLKVGM DVQWDADMSV EEMVNEGVR
101 AYTWEENTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGD KVEVTCAAKG
151 GSEENKSLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGTPEKA
201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELF EK VNALGIGAQQ
251 LGGLTTVL DV KILDYPHTAA SKPIAMIPNC AATRHVEFEL DGSGPVELTP
301 PRVEDWPDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTRDA
351 AHKRLVDMLD KGEELPVDF TNRLIYYVGPV DPGDEIVGP AGPTTATRMD
401 KFTRQMLEQT DLLGMIGKSE RGAATCEAIA DNKAVYLM AV GGAAYLVAKA
451 IKSSKVLAFP ELGMEAIYEF EVKDMPTVTA VDSKGESIHA TAPPQWQAKI
501 GIIPVKs*

```

a050-1/m050-1 98.4% identity in 507 aa overlap

```

a050-1.pep      10      20      30      40      50      60
MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
|||||
m050-1          10      20      30      40      50      60
MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE

a050-1.pep      70      80      90      100     110     120
NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
|||||
m050-1          70      80      90      100     110     120
NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK

a050-1.pep     130     140     150     160     170     180
RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
|||||
m050-1         130     140     150     160     170     180
RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG

a050-1.pep     190     200     210     220     230     240
AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLELFEK
|||||
m050-1         190     200     210     220     230     240
AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLELFEK

a050-1.pep     250     260     270     280     290     300
VNALGIGAQGLGGLTTVLDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVELTP
|||||
m050-1         250     260     270     280     290     300
VNALGIGAQGLGGLTTVLDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVELTP

a050-1.pep     310     320     330     340     350     360
PRVEDWPDLTYS PDNGKRVDVDKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLD
|||||
m050-1         310     320     330     340     350     360
PRVEDWPDLTYS PDNGKRVDVDKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLN

a050-1.pep     370     380     390     400     410     420
KGEELPVDFNRLIYYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE
|||||
m050-1         370     380     390     400     410     420
KGEELPVDFNRLIYYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE

a050-1.pep     430     440     450     460     470     480
RGAATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVVDMPVTVA
|||||
m050-1         430     440     450     460     470     480
RGVATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVVDMPVTVA

a050-1.pep      490      500
VDSKGESIHATAPPQWQAKIGIIPVKSX
|||||
m050-1          490      500
VDSKGESIHATAPRKWQAKIGIIPVESX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 223>:

g052.seq

```

1   ATGGCTTTGG TGGCGGAGGA AACGGAATA TCCGCGCCGT GTTTCAAAGG
51  CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
151 AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
201 GCGGCTTTC CATTCAATTA TATCAGTCGG CGACACGCGG CTCACCCGA
251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
351 CGCCGCCTGA

```

This corresponds to the amino acid sequence &lt;SEQ ID 224; ORF 052.ng&gt;:

g052.pep

1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP  
 51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN  
 101 RLRLETTWSP ACRKVKNA\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 225>:

m052.seq  
 1 ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG  
 51 CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC  
 101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC  
 151 AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC  
 201 GGCGGCTTTC CATTCATTTA TATCAGTCGG CGACACGCGG CTCACTCCGA  
 251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC  
 301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA  
 351 CGCCGCCTGA

This corresponds to the amino acid sequence <SEQ ID 226; ORF 052>:

m052.pep  
 1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP  
 51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN  
 101 RLRLETTWSP ACRKVKNA\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 227>:

a052.seq  
 1 ATGGCTTTGG TCGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG  
 51 CTGAGAGCCG ACAGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC  
 101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCTCCC  
 151 AAGGGATTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC  
 201 GGCGGCTTTC CATTCGTTTA TATCAGTCGG CGACACGTGA CTCACTTCGA  
 251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC  
 301 AGGCTGCGGC TGGAAATCAC ATGGTCGCCC GCCTGCAAAA AGGTGAAAAA  
 351 CGCCGCCTGA

This corresponds to the amino acid sequence <SEQ ID 228; ORF 052.a>:

a052.pep  
 1 MALVAEETEI SAPCFKG\*EP TGDSRLLSTT KSAPMPCANS AKASKSATSP  
 51 KGLDGVSKNS SLVLALTAAF HSFISVGD\* LTSMPLVTM LLIKPTVVPN  
 101 RLRLEITWSP ACKKVKNA\*

m052/a052 95.8% identity over a 119 aa overlap

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
a052	MALVAEETEI SAPCFKGXEP TGDRLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m052.pep	SLVLALTAAFHSFISVGDTRLTPMPNLVTMLLIKPTVVPNRLRLLETTWSPACRKVKNAAX					
a052	SLVLALTAAFHSFISVGDXTLTSMPNLVTMLLIKPTVVPNRLRLLEITWSPACKKVKNAAX					
	70	80	90	100	110	120

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 052 shows 95.8% identity over a 119 aa overlap with a predicted ORF (ORF 052.ng) from *N. gonorrhoeae*:

m052/g052

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					



```

g052      MALVAEETEISAPCFKGCEPTGDSRLSTTKSAPMPCANSASAKSKSATSPKGLDGVSKNS
           10      20      30      40      50      60

           70      80      90      100     110     120
m052.pep  SLVLALTAAFHSFISVGDTWLTSMPLNATMLLIKPTVVPNRLRLEITWSPACKKVKNAAAX
           |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g052      SLVLALTAAFHSFISVGDTRLTTPMPLNLTMLLIKPTVVPNRLRLETTWSPACRKVKNAAX
           70      80      90      100     110     120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 229>:

```

g073.seq
1  ATGTGTATGC CATACGCAAT AAGGGTTTCA GACGGCATCT GCCGCATTTT
51  TCCGCCGATG CCGTCTGAAA CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
151 AGTCCGGGGC GGatacCGGC GGCGAGTTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGGTT GTCCACGGAT TGGTGATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GGAAAAGAGT TCCACGACTT TCCACGCGGC TGCTTGGTCG
301 GCGACTTCAA AACCGATGAC GATGCCGCCG CCGTTTTGCT GTTGCGGGAT
351 AAGCTCCGCC TCGCGATGGT CGGGCAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 230; ORF 073.ng>:

```

g073.pep
1  MCMPYAIRVS DGICRIFFPM PSETRNQRAS ACFKSSIKSP TYSKPTDRRT
51  SPGRIPAASF SSGCILPCVV VHGLVMVERT SPRLAVREKS STTFHAAAWS
101 ATSKPMTMPP PFCCLRISSA CGWSGNPV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 231>:

```

m073.seq
1  ATGTGTATGC CATATAAGAT AAGGGTTTCA GACGGCATCT GCTGTCCAAT
51  GCCGTCTGAA ACACGCAATC AGCGTGCGAG TGCCTGTTTC AAATCGTCAA
101 TCAAATCGCC AACATATTCC AAACCGACCG ACAGGCGCAC CAATCCGGGG
151 CGGATGTTGG CGGCGAGTTT TTCTTCGGGC TGCATCCTGC CGTGCGTGGT
201 TGTCCACGGG TGGGTAATGG TCGAGCGCAC GTCACCGAGG TTGGCGGTGC
251 GGGAAAAGAG TTCCACGCCG TCCACAACCT TCCACGCCGC TTCTTGATCG
301 GCAACTTCAA AGCCGATGAC GATGCCGCCG CCGTTTTGCT GTTGCGGGAT
351 AAGCGCCGCC TGAGGATGGT CGGACAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 232; ORF 073>:

```

m073.pep
1  MCMPYKIRVS DGICCPMPSE TRNQRASACF KSSIKSPTYS KPTDRRTNPG
51  RMLAASFSSG CILPCVVVHG WVMVERTSPR LAVREKSSTP STTFHAASXS
101 ATSKPMTMPP PFCCLRISAA XGWSNPNV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 233>:

```

a073.seq
1  ACGTGTATGT CATATAAGAT AAGGGTTTCA GACGGCATTT GCGGTGTTTT
51  TCCGCCGATG CCGTCTGAA. CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
151 AATCCGGGGC GGATGTTGGC GGCGAGTTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGGTT GTCCACGGAT GGGTAATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GGAGAAAAGT TCGACGCCGT CCACGACTTT CCACGCGGCT
301 GCTTGGTGGC GACTTCAA GCGGATGACG ATGCCGCCGC CGTTTGTCTG
351 TTTGCGGATA AGTCCGCCCT GAGGATGGTC GGGTAATCCG GTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 234; ORF 073.a>:

```

a073.pep
1  TCMSYKIRVS DGICGVFFPM PSEXRNQRAS ACFKSSIKSP TYSKPTDRRT
51  NPGRMLAASF SSGCILPCVV VHGWVVERT SPRLAVREKS STPSTTFHAA
101 AWSATSKPMT MPPPFCLRI SSA*GWSGNP V*

```

m073/a073 92.3% identity over a 130 aa overlap

```

              10      20      30      40      50
m073.pep    MCMPIKIRVSDGICCC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
              || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a073        TCMSYKIRVSDGICGVFPMPSEXRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
              10      20      30      40      50      60

              60      70      80      90      100     110
m073.pep    SSGCILPCVVVHGWMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCLRI
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a073        SSGCILPCVVVHGWMVERTSPRLAVREKSSTPSTTFHAAAWSATSKPMTMPPPFCCLRI
              70      80      90      100     110     120

m073.pep    120      129
              SAAXGWSNPNVX
              |: ||||| |||||
a073        SSAXGWSGNPNVX
              130

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 073 shows 87.0% identity over a 131 aa overlap with a predicted ORF (ORF 073.ng) from *N. gonorrhoeae*:

m073/g073

```

              10      20      30      40      50
m073.pep    MCMPIKIRVSDGICCC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g073        MCMPIAIRVSDGICRIFPPMPSETRNQRASACFKSSIKSPTYSKPTDRRTSPGRIPAASF
              10      20      30      40      50      60

              60      70      80      90      100     110
m073.pep    SSGCILPCVVVHGWMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCLRI
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g073        SSGCILPCVVVHGLVMVERTSPRLAVREKSST---TFHAAAWSATSKPMTMPPPFCCLRI
              70      80      90      100     110

m073.pep    120      129
              SAAXGWSNPNVX
              |: ||||| |||||
g073        SSACGWSGNPNVX
              120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 235>:

g075.seq

```

1   ATGCCGCCTT ACTTCATCAC CCTCTTAACG ATGGAAAATA CAAAAAGCGC
51  GGCGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 CGGCTTCCAA AGCGTTTTTT GCGTTCGCG GCAACGCTGC GTTGCGCTGT
151 GCCGCCAAAG CCAGCGGGGC GGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTT ACGAAATTTT
251 TAAAAAATG  TGTTTGCGGG CTTTGTGAAG GTTTTAGAGA CCGCCTGCCG
301 GGCCTCTTAA ACTTAATCTT CTTTTCGTA GAATCCGAAA ATTACAAATT
351 CCCCCTAT  CTCTTCCAAT GCCGAGCTAA AAGCGTCTC ATAGCTGTCA
401 TATTACCGG CTGA

```

This corresponds to the amino acid sequence <SEQ ID 236; ORF 075.ng>:

g075.pep

```

1   MPPYFITLLT MENTKSAKT PTTIQPASIP SAFAASKAFF AVSGNAAFAC
51  AAKASGAAVT TASFAPYLQ VLINFMISF TKFLKKCVCG LCEGFRDRLP
101 GLNLIFFFV  ESENYKPAY LFQCRKSVF IAVIFTG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 237>:

```
m075.seq
  1 ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAATAA CAAAAAGCGC
 51 GGCAGAAATG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 CGGCTTCCAA AGCGTTTTTT GCCGTATCGG GCAACGTTGC ATTTGCATGT
151 GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTC AAAAAGTGTT
251 TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301 TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351 CTTCCAAACC TCGGTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401 TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 238; ORF 075>:

```
m075.pep
  1 MPSYFITLLT MENTKSAAKM PTTIQPASIP SAFAASKAFF AVSGNVAFAC
 51 AAKARGAAVT TASFAPYLRQ VLINFMIFSF KKCLAVMDGA FRRPPNIRK
101 SVFQKSEYDK FVLVADFFQT CVNRFFEVEE IIGIGD*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 075 shows 65.7% identity over a 137 aa overlap with a predicted ORF (ORF 075.ng) from *N. gonorrhoeae*:

```
m075/g075

      10      20      30      40      50      60
m075.pep MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
          |||||||
g075      MPPYFITLLTMENTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNAAAFACAAKASGAAVT
          |||||||

      70      80      90      100     110
m075.pep TASFAPYLRQVLINFMIFSF----KKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVAD
          |||||||
g075      TASFAPYLRQVLINFMIFSFTKFLKKVCVGLCEGFRDRLPGLLNLIFFVFSESENYKFPAY
          |||:  :  :|  |::  :  :|  |::  :  :|

      70      80      90      100     110     120
m075.pep
          FFQTCVNRFFEVEEIIIGIGDX
          :||  ::  |  :  :  |
g075      LFQCRAKSVFIAVIFTGX
          130
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 239>:

```
a075.seq
  1 ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAATAA CAAAAAGCGC
 51 GGCAGAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 CGGCTTCCAA AGCGTTTTTT GCTGTATCGG GCAACGTTGC ATTTGCATGT
151 GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTC AAAAAGTGTT
251 TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301 TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351 CTTCCAAACC TCGGTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401 TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 240; ORF 075.a>:

```
a075.pep
  1 MPSYFITLLT MEKTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNVAFAC
 51 AAKARGAAVT TASFAPYLRQ VLINFMIFSF KKCLAVMDGA FRRPPNIRK
101 SVFQKSEYDK FVLVADFFQT CVNRFFEVEE IIGIGD*
```

m075/a075 98.5% identity over a 136 aa overlap

```

      10      20      30      40      50      60
m075.pep MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFAAKARGAAVT
          |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a075      MPSYFITLLTMEKTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNVAFAAKARGAAVT
          10      20      30      40      50      60

      70      80      90     100     110     120
m075.pep TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a075      TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
          70      80      90     100     110     120

      130
m075.pep  CVNRFFEVEIIGIGDX
          |||||
a075      CVNRFFEVEIIGIGDX
          130

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 241>:

```

g080.seq
1   ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAAGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA TACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGGTTGCA CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAA TGTCGTTTTG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACgaGA TGAAACGCCT CCgGCTTTTT
601 ACcgAAGCGT GGCAGCATCT gttgcGTAAG AATAAAAATC GGTATCCTA
651 TGTGGATATG Aggtataagg acggatttTC agtcccccat gctCCCGACG
701 GTTTACCCGA AAAAGAATcc gAAGAATatt gggaaacaggt ttgggacata
751 ttacggcctg gcgtcggaaa cggttcgcacg caaatttcaa tcagttataA
801 GGGCAGacga acaatggaac AGcagtaa

```

This corresponds to the amino acid sequence <SEQ ID 242; ORF 080.ng>:

```

g080.pep
1   MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLVPK QVSLKGNLVY
51  SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPIWIASVM VRRRFPDIVE
101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPYMPVFRGA EGTSAEMLRR
151 YDEFSTVLAK QGLGIKEMTY TARSAWNVVL DNGITVRLGR ENEMKRLRLF
201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVPH APDGLPEKES EEYWEQVWDI
251 LRPVGNGST QISISYKRR TMEQQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 243>:

```

m080.seq
1   ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGACATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGGTCCGG CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTTG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT

```

601 ACCGAAGCGT GGCAGCATCT GTTGCCTAAA AATAAAAATC GGTATCCTA  
 651 TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTTCCGACG  
 701 GTTTACCCGA AAAAGAATCC GAAGAATAG

This corresponds to the amino acid sequence <SEQ ID 2441; ORF 080>:

m080.pep  
 1 MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYN SNHLPVK QVSLKGNLVY  
 51 SDKKTLGSLA KEYIHGNILR TDINGAQEAY RRY PWIASVM VRRRFPDTVE  
 101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EG TSAEMLRR  
 151 YDEFSTVLAK QGLGIKEMTY TARS AWIVVL DNGITVRLGR ENEMKRLRLF  
 201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY ASDGLPEKES EE\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 080 shows 97.9% identity over a 242 aa overlap with a predicted ORF (ORF 080.ng) from *N. gonorrhoeae*:

m080/g080

	10	20	30	40	50	60
m080.pep	MWDNAEAMERLTRWLLVMMA MLLAASGLVWFYNSNHLVPVKQVSLKGNLVYSDKKTLGSLA					
080	MWDNAEAMERLTRWLLVMMA MLLAASGLVWFYNSNHLVPVKQVSLKGNLVYSDKKALGSLA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m080.pep	KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG					
080	KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG					
	70	80	90	100	110	120
	130	140	150	160	170	180
m080.pep	EGNVFEARLDRPGMPVFRGAEG TSAEMLRRYDEFSTVLAKQGLGIKEMTYTARS AWIVVL					
080	EGNVFEARLDRPGMPVFRGAEG TSAEMLRRYDEFSTVLAKQGLGIKEMTYTARS AWNVVL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m080.pep	DNGITVRLGRENEMKRLRLFTEAWQHLLRK NKNRLSYVDMRYKDGFSVRYASDGLPEKES					
080	DNGITVRLGRENEMKRLRLFTEAWQHLLRK NKNRLSYVDMRYKDGFSVPHAPDGLPEKES					
	190	200	210	220	230	240
m080.pep	EEX					
080	EEYWEQVWDILRPGVGN GSTQISISYKGRRTMEQQX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 245>:

a080.seq  
 1 ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT  
 51 CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT  
 101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTAGTTTAT  
 151 TCCGATAAGA AAGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA  
 201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC  
 251 CGTGGATTGC GTCGGTCATG GTGCGCGGCC GTTTCCCGA CACGGTTGAG  
 301 GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT  
 351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGTTTGGAC AGACCCGGAA  
 401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT  
 451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA  
 501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTGTTTTTG GACAACGGCA

551 TCACCGCTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT  
 601 ACCGAAGCGT GGCAACATCT GTTGCCTAAA AATAAAAATC GGTTATCCTA  
 651 TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTCCCGACG  
 701 GTTTACCCGA AAAAGAATCC GAAGAATAG

This corresponds to the amino acid sequence <SEQ ID 246; ORF 080.a>:

## a080.pep

1 MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYN SNHLPVK QVSLKGNLVY  
 51 SDKKALGSLA KEYIHGNILR TDINGAQEAY RYPWIASVM VRRRFPDIVE  
 101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EG TSAEMLRR  
 151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF  
 201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY APDGLPEKES EE\*

m080/a080 99.2% identity over a 242 aa overlap

	10	20	30	40	50	60
m080.pep	MWDNAEAMERL	TRWLLVMMA	MLLAASGLVW	FYN SNHLPVK	QVSLKGNLVY	SDKKTLGSLA
a080	MWDNAEAMERL	TRWLLVMMA	MLLAASGLVW	FYN SNHLPVK	QVSLKGNLVY	SDKKALGSLA
	70	80	90	100	110	120
m080.pep	KEYIHGNILR	TDINGAQEAY	RRYPWIASVM	VRRRFPDIVE	VVLTERKPVA	RWGDHALVDG
a080	KEYIHGNILR	TDINGAQEAY	RRYPWIASVM	VRRRFPDIVE	VVLTERKPVA	RWGDHALVDG
	130	140	150	160	170	180
m080.pep	EGNVFEARLD	RPGMPVFRGA	EG TSAEMLRR	YDEFSTVLAK	QGLGIKEMTY	TARSAWIVVL
a080	EGNVFEARLD	RPGMPVFRGA	EG TSAEMLRR	YDEFSTVLAK	QGLGIKEMTY	TARSAWIVVL
	190	200	210	220	230	240
m080.pep	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLRK	NKNRLSYVDM	RYKDGFSVRY	ASDGLPEKES
a080	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLRK	NKNRLSYVDM	RYKDGFSVRY	APDGLPEKES
m080.pep	EEX					
a080	EEX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 247>:

## g081.seq

1 ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT  
 51 GCCGTCTGAA AACAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGATA  
 101 TTCGGAAGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGACGCG  
 151 CATGATTTTG TTGGAGGCGT ATTGTCTGCG GCGCGGCGCG CGGTTGTGGT  
 201 TTCGCGCGAA GATTGCGCGG CTTTGGGCGG CGCGTTGAAA GTCGATGACA  
 251 CGCTTGCCGC GTTGCAAACG TTGGCGAAGG CGTGGCGCGA TAATGTGAAC  
 301 CCGTTTGTGT TCGGCATTAC CGGTTTCGGC GGCAAGACGA CGGTGAAGGA  
 351 GATGCTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATGAT GCCGTTTCGG  
 401 CGACGGCAGG CAACTTCAAC AACCACAtcg gaTTGCCGCT GACTTTATTG  
 451 AAATtaacg aAAAAACCCG CTATGCCGTG ATTGAAATGG GCATGAACCA  
 501 TTTTGGcgaa ctggcggtTt taacgcaaaT CGCCAAACCC GATGCCGCTT  
 551 TGGtcaACAA CGCCCTGCGC GCCCATGTCG GATGCGGTTt cgacggagtg  
 601 GCGCATATTG CCAAAGcgaa aagcGAGATT TatgcagGct tATGTTTCAAG  
 651 CGGCATGGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA  
 701 CGGCAACGTT TAATTTGAAT ACGTGCACTT TCGGCGTCGA TAGCGCGGAT  
 751 GTCCGCGCGG AAAATATCGT GCTGAAACCT TTGTCGTGCG AATTTGATT

```

801 GGTGTGCGGC GACGAGCGCA CTGCCGTGGT GCTGCCTGTT CCCGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCCGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG CAAGGCTTCA GCAACATCAA
951 AGGCCGTCTG AACGTCAAAG CCGGCATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAATCCC GACAGTATGA AAGCCGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG CATTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAGGAc gaAGCCGCCG CCATGCACGC CGAagtcgGC GCGTACGCCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 GCGGcggaaA AATTTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTGTTGAGCC ACGATTGGCC CGAACGCGCC ACCGTGTTGG
1301 TGAAAGGTTT GCGCTTTATG CAGAtggaAG AAGTGGTCGA GGCATTGGAG
1351 GATAAGTga

```

This corresponds to the amino acid sequence <SEQ ID 248; ORF 081.ng>:

```

g081.pep
  1 MKPLDLNFIC QALKLPMPSE NKPVSRIVTD SRDIREGDVF FALAGGRFDA
 51 HDFVGGVLSA GAAAVVVSRE DCAALGGALK VDDTLAALQT LAKAWRDNVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVSATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFE LAVLTQIAKP DAALVNNALR AHVGCDFDGV
201 GDIAKAKSEI YAGLCSDGMA LIPQEDANMA VFKTATFNLN TCTFGVDSGD
251 VRAENIVLKP LSCEFDLVCG DERTAVVLPV PGRHNVHNA AAAAAALAAG
301 LSLNDVAEGL QGFSNIGKRL NVKAGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGDMGELGED EAAAMHAEVG AYARDQIEA AYFVGDNVSE
401 AAEKFGADGL WFAAKDPLIQ VLSHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 249>:

```

m081.seq
  1 ATGAAACCAC TGGACCTAAA TTTTCATCTGC CAAGCCCTCA AGCTTCCGAT
 51 GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGGCGAGCG GTTGACGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCT GGTGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGTGCTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTGCCGC ATTGCAAACG CTGGCAAAGG CGTGGCGTGA AAATGTGAAT
301 CCGTTTGCTG TCGGCATTAC CGGTTCGGGC GGCAAGACGA CGGTGAAGGA
351 AATGCTGGCT GCGGTATTGC GCCgCCGTTT CGGCGATGAT GCCGTGTTGG
401 CGACGCGAGG CAACCTTCAAC AACCATATCG GATTGCCGCT GACTTTGTTG
451 AAGTTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTCGCGCAA CTGGCGGTTT TAACGCaMAT CGCCAAACCA AATGCCGCAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTCG GCTGCGGTTT CGACGGAGTG
601 GGCGATATTG CCAAAGCGAA AAGCGAGATT TACCAAGGTT TATGTTTCTG
651 CGGCATTGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGCT TAATTTGAAT ACGGCACTT TCGGCATCGA TAGCGGCGAT
751 GTTCACGCGG AAAATATTGT GCTGAAACCG TTGTCGTGCG AATTTGATTT
801 GGTGTGCGGC GATGAGCGCG CCGCCGTGGT GCTGCCTGTT CCCGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCGGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAACCCT GACAGCATGA AAGCTGCGAT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAACTG GGCGAGGACG AAGCCGCCGC TATGCACGCC GAAGTCGGCG
1151 CGTATGCCCG CGACCAAGGC ATCGAAGCGG CTTATTTTGT CGGCGACAAC
1201 AGCGTCGAAG CGGCGGAAAA ATTTGGCGCG GACGGTTTGT GGTTCGCCGC
1251 CAAAGACCCG TTGATTCAAG TGTTGCGCCA CGATTTGCCC GAACGCGCCA
1301 CCGTGTGGT GAAAGGTTTC CGCTTTATGC AGATGGAAGA AGTGGTCGAG
1351 GCATTGGAGG ATAAGTGA

```

This corresponds to the amino acid sequence <SEQ ID 250; ORF 081>:

```

m081.pep
  1 MKPLDLNFIC QALKLPMPSE SKPVSRIVTD SRDIRAGDVF FALAGERFDA
 51 HDFVEDVLA A GAAAVVVSRE DCAAMDGALK VDDTLAALQT LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFE LAVLTXIAKP NAALVNNAMR AHVGCDFDGV

```

201 GDIKAKSEI YQGLCSDGIA LIQEDANMA VFKTATLNLN TRTFGIDSGD  
251 VHAENIVLKP LSCEFDLVCG DERAADVLPV PGRHNVHNAA AAAALALAAG  
301 LSLNDVAEGL KGFSNIKGRL NVKSGIKGAT LIDDTYNANP DSMKAAIDVL  
351 ARMPAPRIFV MGDMGELGEL GEDEAAAMHA EVGAYARDQG IEAAYFVGDN  
401 SVEAAEKFGA DGLWFAAKDP LIQVLRHDLR ERATVLVKGS RFMQMEEVVE  
451 ALEDK\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 081 shows 94.1% identity over a 455 aa overlap with a predicted ORF (ORF 081.ng)  
from *N. gonorrhoeae*:

m081/g081

m081.pep	10	20	30	40	50	60
	MKPLDLNFICQALKLPMPSESKPVSRIVTDSRDIRAGDVFFALAGERFDAHDFVEDVLAA					
g081	MKPLDLNFICQALKLPMPSENKPVSRIVTDSRDIREGDVFFALAGGRFDAHDFVGGVLSA					
	10	20	30	40	50	60
m081.pep	70	80	90	100	110	120
	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
g081	GAAAVVVSREDCAALGGALKVDDTLAALQTLAKAWRDNVNPFVFGITGSGGKTTVKEMLA					
	70	80	90	100	110	120
m081.pep	130	140	150	160	170	180
	AVLRRRFGDDAVLATAGNFNNHIGLPLTLLKLNKHYAVIEMGMNHFGEALVLTIXIAKP					
g081	AVLRRRFGDDAVSATAGNFNNHIGLPLTLLKLNKHYAVIEMGMNHFGEALVLTQIAKP					
	130	140	150	160	170	180
m081.pep	190	200	210	220	230	240
	NAALVNNAMRAHVGCDFGVGDIKAKSEIYQGLCSDGIALIPQEDANMAVFKTATLNLN					
g081	DAALVNNALRAHVGCDFGVGDIKAKSEIYAGLCSDGMALIPQEDANMAVFKTATFNLN					
	190	200	210	220	230	240
m081.pep	250	260	270	280	290	300
	TRTFGIDSGDVHAENIVLKLPLSCEFDLVCGDERAADVLPVPGRHNVHNAAAAALALAAG					
g081	TCTFGVDSGDVRAENIVLKLPLSCEFDLVCGDERTAVVLPVPGRHNVHNAAAAALALAAG					
	250	260	270	280	290	300
m081.pep	310	320	330	340	350	360
	LSLNDVAEGLKGFSNIKGRLNVKSGIKGATLIDDTYNANPDSMKAAIDVLARMPAPRIFV					
g081	LSLNDVAEGLQGFSNIKGRLNVKAGIKGATLIDDTYNANPDSMKAAVDVLARMPAPRIFV					
	310	320	330	340	350	360
m081.pep	370	380	390	400	410	420
	MGDMGELGELGEDEAAAMHAEVGAYARDQGEAAYFVGDNSEAAEKFGADGLWFAAKDP					
g081	MGDMGELGE---DEAAAMHAEVGAYARDQGEAAYFVGDNSEAAEKFGADGLWFAAKDP					
	370	380	390	400	410	
m081.pep	430	440	450			
	LIQVLRHDLPERATVLVKGSRFMQMEEVVEALEDKX					
g081	LIQVLSHDLPERATVLVKGSRFMQMEEVVEALEDKX					
	420	430	440	450		



The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 251>:

## a081.seq

```

1   ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51  GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGATGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCG GGTGCGGGCG CGGTGTGGT
201 TTCGCGCGAA GATTGCGTTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTACCGC GTTGCAAATG TTGGCGAAGG CGTGGCGCGA GAATGTGAAC
301 CCGTTTGTGT TCGGTATTAC CGGCTCGGGC GGCAAGACGA CGGTGAAGGA
351 AATGTTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATAAT GCCGTTTGG
401 CGACGGCAGG CAACTTCAAC AACCACATCG GATTGCCGTT GACTTTGTTG
451 AAATTAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GTATGAACCA
501 TTTTGGCGAA CTGGCGGTTT TGACACAAAT CGCCAAACCC GATGCCGCAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTCG GCTGCGGTTT CGACGGAGTG
601 GCGGATATTG CCAAAGCGAA AAGCGAGATT TATCAAGGCT TATGTTTCTG
651 CCGCATGGCG CTGATTCCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 TTCCACGCGG AAAATATCGT GCTGAAACCG TTGTCGTGCG AATTTGATT
801 GGTGTGCGGC AACGAGTGGC CAGCCGTGGT TCTGCCCGTT CCCGCCGCC
851 ACAATGTCCA CAACGCCGCC GCCGCCGCCG CGCTGTCTTT GGCTGCAGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGCTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAACCCCT GACAGCATGA AAGCTGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGTGAGGAC GAAGCCGCCG CCATGCACGC CGAAGTCGGC GCGTACGCCC
1151 GCGACCAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 GCGGCGGAAA AATTTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTGTTGCGCC ACGATTGACC CGAACGCGCC ACCGTGTTGG
1301 TGAAAGGTTT GCGCTTTATG CAGATGGAAG AAGTGGTCTGA GGCATTGGAG
1351 GATAAGTGA

```

This corresponds to the amino acid sequence <SEQ ID 252; ORF 081.a>:

## a081.pep

```

1   MKPLDLNFIC QALKLMPSE SKPVSRIVTD SRDIRAGDVF FALAGGRFDA
51  HDFVEDVLAA GAAAVVVSRE DCVAMDGALK VDDTLTALQM LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDN AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTQIAKP DAALVNAMR AHVCGFDGV
201 GDIKAKSEI YQGLCSGMA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
251 VHAENIVLKP LSCEFDLVCG NECAAVVLPV PGRHNVHNA AAAALS LAAG
301 LSLNDVAEGL KGFSNIKGR LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGD MGELGED EAAAMHAEVG AYARDQIEA AYFVGDN SVE
401 AAEKFGADGL WFAAKDPLIQ VLRHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

```

m081/a081 96.7% identity over a 455 aa overlap

	10	20	30	40	50	60
m081.pep	MKPLDLNFICQALKLMPSESKPVSRIVTDSDIRAGDVFFALAGERFDAHDFVEDVLAA					
a081	MKPLDLNFICQALKLMPSESKPVSRIVTDSDIRAGDVFFALAGGRFDAHDFVEDVLAA					
	10	20	30	40	50	60
m081.pep	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
a081	GAAAVVVSREDCVAMDGALKVDDTLALQMLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
	70	80	90	100	110	120
m081.pep	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
a081	GAAAVVVSREDCVAMDGALKVDDTLALQMLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
	130	140	150	160	170	180
m081.pep	AVLRRRFGDDAVLATAGNFNNHIGLPLTLLKLNKHRYAVIEMGMNHFGE LAVLTIXIAKP					
a081	AVLRRRFGDNAVLATAGNFNNHIGLPLTLLKLNKHRYAVIEMGMNHFGE LAVLTQIAKP					

	130	140	150	160	170	180
m081.pep	190	200	210	220	230	240
	NAALVNNAMRAHVGC GFDGVGDI AKAKSEIYQGLCSDGIALIPQEDANMAVFKTATLNLN					
a081	:					
	190	200	210	220	230	240
	DAALVNNAMRAHVGC GFDGVGDI AKAKSEIYQGLCSDGMALIPQEDANMAVFKTATLNLN					
m081.pep	250	260	270	280	290	300
	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGDERAAVVLVPVGRHNVHNAALALAAG					
a081						
	250	260	270	280	290	300
	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGNECAAVVLVPVGRHNVHNAALSLAAG					
m081.pep	310	320	330	340	350	360
	LSLNDVAEGLKGFSNIKGRNLNVKSGIKGATLIDDTYNANPDSMKAAIDVLARMPAPRIFV					
a081						
	310	320	330	340	350	360
	LSLNDVAEGLKGFSNIKGRNLNVKSGIKGATLIDDTYNANPDSMKAAVDVLARMPAPRIFV					
m081.pep	370	380	390	400	410	420
	MGDMGELGELGEDEAAAMHAEVGAAYARDQIEAAYFVGDNVSVEAAEKFGADGLWFAAKDP					
a081						
	370	380	390	400	410	420
	MGDMGELGE---DEAAAMHAEVGAAYARDQIEAAYFVGDNVSVEAAEKFGADGLWFAAKDP					
m081.pep	430	440	450			
	LIQVLRHDLPERATVLVKGSRFMQE EVVEALEDKX					
a081						
	420	430	440	450		
	LIQVLRHDLPERATVLVKGSRFMQE EVVEALEDKX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 253>:

g082.seq

1	aTGTGGTTGT	TGAAGTTGCC	TGCCGTCGCC	GAAACGGCAT	CATCGCCGAA
51	ACGGCGGCGC	AATACCGCAG	CCAGCATCTC	CTTACCGGTC	GTCTTGCCGC
101	CCGAACCGGT	AATGCCGAAC	ACAAACGGGT	TCACATTATC	GCGCCACGCC
151	TTGCGCAACG	TTTGCAACGC	GGCAAGCGTG	TCATCGACTT	TCAACGCGCC
201	GCCCAAAGCC	GCGCAATCTT	CGCGCGAAAC	CACAACCGCC	GCCGCGCCCG
251	CAGACAATAC	GCCTCCAACA	AAATCATGCG	CGTCAAACCG	CCCGCCCGCC
301	AATGCGAAAA	ACACATCGCC	TTCCCGAATA	TCGCGGCTGT	CGGTTACGAT
351	GCGCGACACG	GGTTTGTTTT	CAGACGGCAT	CGGAAGCTTG	AGGGCTTGGC
401	AGATGAAATT	TAGGTCCAGT	GGTTTCATAT	TTGCTTTCGT	TAATATTCGG
451	GCGGCGGACA	CATCGGTAGC	GGCTGATTTT	TTATCGCCT	GTTTGCTGT
501	GGTAAACAC	AGATTATTTT	CCCATTCTCA	TTGCGCATT	TTTCTGTACG
551	TATCATTTTT	TAGACGTATT	TTAGCCGAT	TTGCCTTTTC	CCGCATACCA
601	CGGCGCGGGG	TCGTCCGACT	GTCTGTCGAT	AAAGGCAAGG	TTATTGCCTT
651	CGCCCGGCAC	ATCGGGGACA	TTCCCCCAA	AATCATAGCC	GTCATCGGGC
701	AACTCGTCGG	TTTCGATACC	CGTCCAACG	CCGAATCCG	GTAA

This corresponds to the amino acid sequence <SEQ ID 254; ORF 082.ng>:

g082.pep

1	MWLLKLP	ETASSPKRRR	NTAASISFTV	VLPPEPVMPN	TNGFTLSRHA
51	FANVCNAASV	SSTFNAPPKA	AQSSRETTTA	AAPADNTPPT	KSCASNRP
101	NAKNTSPSRI	SRLSVTMRDT	GLFSDGIGSL	RAWQMKFRSS	GFIFAFVNIR
151	AADTSVAADF	FIACFAVVKH	RLFSHSHSAF	FLYVSFFERRI	FSRFAFSRIP
201	RRGVVGLSVD	KGKVIAFARH	IGDIPPKIIA	VIGQLVGFD	RPTAES*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 255>:

m082.seq

1	ATGnnGTTGT	TGAAGTTGCC	TGCCGTCGCC	AACACGGCAT	CATCGCCGAA
51	ACGGcGGCGC	AATACCGCAG	CCAGCATTTT	CTTACCGGTC	GTCTTGCCGC
101	CCGAACCGGT	AATGCCGAAC	ACAAACGGAT	TCACATTTTC	ACGCCACGCC
151	TTTGCCAGCG	TTTGCAATGC	GGCAAGCGTG	TCATCGACTT	TCAACGCGCC

201	ATCCATTGCA	GCACAATCTT	CGCGCGAAAC	CACAACCGCC	GCCGCAACCAG
251	CAGCCAATAC	GTCTTCAACA	AAATCATGCG	CGTCAAACCG	CTCGCCCGCC
301	AATGCGAAAA	ACACATCGCC	CGCGCGGATG	TCGCGGCTGT	CGGTTACGAT
351	GCGCGACACG	GGTTTGCTTT	CAGACGGCAT	CGGAAGCTTG	AGGGCTTGGC
401	AGATGAAATT	TAGGTCCAGT	GGTTTCATAT	TTACTTTCGT	TAATATTTCGG
451	GCGGCGGACA	CATCGGTAGC	GGCTGATTTT	TTTATCGCCT	GTTTTGCTGT
501	GGTAAAACAC	AGATTATTTT	CCCATTCTCA	TTGGcsATTT	TTTCTGTACG
551	TATCATTTTT	TAGACGTATT	TTTAGTCGAT	TTGCGTTTTT	CCGATACCA
601	CGGCGCGGGG	TCGTCGGGCA	GTCCGTCGAT	AAAGGCAAGG	TATTGCGCTT
651	CGCCCTGCAC	ATCGGGAACA	TTCCCCAAA	AATCATAGCC	GTCATCGGGC
701	AACTCGTCGG	TTTCGATACC	CGTCCAAC TG	CCGAATCCGC	GTAA

This corresponds to the amino acid sequence <SEQ ID 256; ORF 082>:

```
m082.pep
1  MXLLKLPAVA NTASSPKRRR NTAASISFTV VLPPEPVMPN TNGFTFSRHA
51  FASVCNAASV SSTFNAPSIA AQSSRETTTA AAPAANTSST KSCASNRSRA
101 NAKNTSPARM SRLSVTMRDT GLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151 AADTSVAADF FIACFAVVKH RLFSHSHSXF FLYVSFFERRI FSRFAFSRIP
201 RRGVVQSQSD KGVIAFALH IGNIPPKIIA VIGQLVGFD TPTAES*
```

Computer analysis of this amino acid sequence gave the following results:

### Homology with a predicted ORF from *N. gonorrhoeae*

ORF 082 shows 92.7% identity over a 247 aa overlap with a predicted ORF (ORF 082.ng) from *N. gonorrhoeae*:

```
m082/g082
```

	10	20	30	40	50	60
m082.pep	MXLLKLP	AVANTASSPKRRRNTAASISFTVVLPPPEVMPNTNGFTFSRHAFASVCNAASV				
		:	:	:	:	
g082	MWLLKLP	AVAETASSPKRRRNTAASISFTVVLPPPEVMPNTNGFTLSRHAFANVCNAASV				
		:	:	:	:	
	70	80	90	100	110	120
m082.pep	SSTFNAPSIAAQSSRETTTAAAPANTSSTKSCASNRS	PANAKNTSPARMSRLSVTMRDT				
						:
g082	SSTFNAPPKAAQSSRETTTAAAPADNTPPTKSCASNRP	PANAKNTSPSRISRLSVTMRDT				
						:
	70	80	90	100	110	120
	130	140	150	160	170	180
m082.pep	GLLSDGIGSLRAWQMFRSSGFIFTFVNIRAADTSVAADFFIACFAVVVKHRLF	SHSHSXF				
	:		:	:	:	
g082	GLFSDGIGSLRAWQMFRSSGFIFAFVNIRAADTSVAADFFIACFAVVVKHRLF	SHSHSAF				
	130	140	150	160	170	180
	190	200	210	220	230	240
m082.pep	FLYVSFFRRIFSRFAFSRI	PRRGVVGQSVDKGKVIAFALHIGHNIPPKII	IAVIGQLVGFD			
g082	FLYVSFFRRIFSRFAFSRI	PRRGVGLSVDKGKVIAFARHIGDIPPKII	IAVIGQLVGFD			
	190	200	210	220	230	240
m082.pep	RPTAESAX					
q082	RPTAESAX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 257>:

**a082.seq**

```
1  ATGTGTTGT  TGAAGTTGCC  TGCCGTCGCC  AAAACGGCAT  TATCGCCGAA
51 ACGGCGGCGC  AATACCGCAG  CCAACATTTC  CTTACCGTTC  GTCTTGCCGC
```

**a082.ppt**

1	2	MWLLKLP	AVA	KTALSPK	RRR	NTAANIS	FTV	VLPPE	VP	IPN	TNGFT	SRHA
51		FANICNA	VSV	SSTFN	APSIA	TQSSRE	TTTA	AAPAAN	TSS		KSCASN	RPPA
101		NAKNTSP	ARM	SRLSV	TMRDT	GLLSD	GIGSL	RAWQM	KFRSS		GFITF	VNIR
151		AADTS	VAADE	FIACF	AVVKH	RLFSD	SHSAF	FLYVS	FFRR		FSRFA	SRIP
201		RRGVV	GQSD	KGKVI	AFAH	IGNIP	PKIIA	VIGQL	VGFD		RTAES	*

**m082/a082** 95.5% identity over a 247 aa overlap

		10	20	30	40	50	60
m082.pep		MXLLKLPAVANTASSPKRRRNTAASISFTVVLPPPEVMPNTNGFTFSRHAFASVCNAASV					
a082		:                 :                 :                 :                 :					
		10	20	30	40	50	60
		MWLLKLPAVAKTALSPKRRRNTAANISFTVVLPPPEVIPNTNGFTFSRHAFANICNAVSV					
		70	80	90	100	110	120
m082.pep		SSTFNAPSIAAQSSRETTTAAAPAANTSSTKSCASNRSPANAKNTSPARMSRLSVTMRDT					
a082		:                 :                 :                 :					
		SSTFNAPSIATQSSRETTTAAAPAANTSSTKSCASNRSPANAKNTSPARMSRLSVTMRDT					
		70	80	90	100	110	120
		130	140	150	160	170	180
m082.pep		GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF					
a082		:                 :                 :                 :					
		GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF					
		130	140	150	160	170	180
		190	200	210	220	230	240
m082.pep		FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT					
a082		:                 :                 :                 :					
		FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT					
		190	200	210	220	230	240
m082.pep		RPTAESAX					
a082							
		RPTAESAX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 259>:

g084.seq

1	ATGAAacaAT	CCGcccgaat	aAAAAATATG	GATCAGACAT	TAAAAAATAc
51	attgggcatt	tGCGCGcttt	tagcctTTTG	TTTg'gcgcG	gccaTCGCAT
101	CAGGTTATCA	CTTGGAATAT	GAATACGGCT	ACCGTTATTc	TGCCGTGGGC
151	GCTTTGGCTT	CGGTTGTATT	TTTATTATTA	TTGGCACGCG	GCTTCCCGCG
201	CGTTTCTTCA	GTGTTTTAC	TGATTTACGT	CGGCACAACC	GCCCTATATT
251	TGCCGCTCGG	CTGGCTTGAT	GGTGGCGCTT	CTTATCAGAT	AGTCGGTTCG
301	ATATTGAAAA	GCAATCCTGC	CGAGGCGCGT	GAAATTGTGC	GCAATCTTCC
351	CGGGTCGCTT	TATTTTGTGC	AGGCATTATT	TTTCATTTTT	GCGTTGACAG

```

401 TTTGGAATA TTGTGTATCT GTGGGGGTAT TTGCTGACGT AAAAACTAT
451 AAACGTCGCA GCAAAATATG GCTGACCATA TTATTGACTT TGATTTTGTC
501 CTGCGCGGTG ATGGAGAAAA TCGccggcga taaAGATTGG CGAGaacctg
551 atgccggcct gttgttgaat ATTTcgacc tgtattaCga cttggctttc
601 cgcgccggca cAATATGCCG CCAAGCGCGC CCACattttg gaagCagcaa
651 aaaaagcgtC AACATGGCAT atccgccaac ttgcgcccaa gTataa

```

This corresponds to the amino acid sequence <SEQ ID 260; ORF 084.ng>:

```

g084.pep
  1 MKQSARIKMN DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
 51 ALASVVFLLL LARGFPRVSS VLLLIYVGTI ALYLPVGWLY GAPSYQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS VGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MEKIAGDKDW REPDAGLLLN IFDLYYDLAF
201 RAGTICRQAR PHFGSSKSV NMAYPPTCAQ V*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 261>:

```

m084.seq
  1 ATGAAACAAT CCGCCcGAAT AAAa.ATATG AATCAGACAT TACTTTATAC
 51 ATTGGGCATT TGCGCGCTTT TAACCTTTnn nnnnnnnnnn nnnnnnnnnn
101 nnnnnTATCA CCCnGAATAT GAATACGGCT ACCGTATTTC TGCCGTGGGT
151 GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GTTCCCCGCG
201 CGTTTCTTCA GTTGTTTTAC TGATTACGT CGGCACAACC GCCCTATATT
251 TGCCGGTCGG CTGGCTGTAT GGTGCGCGT CTTATCAGAT AGTCGGTTCG
301 ATATTGGAAG GCAATCCTGC CGAGGCGCGT GAATTGTGCG GCAATCTTCC
351 CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCATTTT GGCTTGACAG
401 TTTGGAATA TTGTGTATCG GGGGGGGTAT TTGCTGACGT AAAAACTAT
451 AAACGCCGCA GCAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTC
501 CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTG CGAGAACCTG
551 ATGCCGCCTT GTTGTGAAT ATTTTCGACC TGTATTACGA TTTGGCT.TC
601 CGCGCCGCA CAATATGCCG CCAAGCGCGC CCACATTTTG GAAGCAGCAA
651 AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA

```

This corresponds to the amino acid sequence <SEQ ID 262; ORF 084>:

```

m084.pep
  1 MKQSARIKXM NQTLTYTLGI CALLTFXXXX XXXXXYHPEY EYGYRYSAVG
 51 ALASVVFLLL LARGFPRVSS VLLLIYVGTI ALYLPVGWLY GAPSYQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS GGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDAGLLLN IFDLYYDLAX
201 RAGTICRQAR PHFGSSKSV NMAYPSCAQ V*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 084 shows 90.5% identity over a 231 aa overlap with a predicted ORF (ORF 084.ng) from *N. gonorrhoeae*:

m084/g084

```

          10      20      30      40      50
m084.pep  MKQSARIKXMNQTLTYTLGICALLTF-----YHPEY EYGYRYSAVGALASVVFLLL
          |||||  ||:|  |||||  ||:|  |||||  |||||  |||||  |||||  |||||
g084      MKQSARIKMNQTLKNTLGIALLAFCFGAAIASGYHLEYEYGYRYSAVGALASVVFLLL
          10      20      30      40      50      60

          60      70      80      90     100     110
m084.pep  LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g084      LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
          70      80      90     100     110     120

          120     130     140     150     160     170
m084.pep  YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g084      YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMEKIAGDKDW
          130     140     150     160     170     180

```

	180	190	200	210	220
m084.pep	REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVMAYPSCCAQVX				
g084	REPDAGLLLNIFDLYYDLAFRAGTICRQARPHFGSSKKSVMAYPPTCAQVX				
	190	200	210	220	230

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 263>:

a084.seq

```

1  ATGAAACAAT CCGCCCGAAT AAAAAATATG GATCAGACAT TAAAAAATAC
51  ATTGGGCATT TGC GCGCTTT TAGCCTTTTG TTTTGGCGCG GCCATCGCAT
101 CAGGTATATCA CTTGGAATAT GAATACGGCT ACCGTTATTC TGCCGTGGGT
151 GCTTTGGCTT CGGTGTGATT TTTATTATTA TTGGCACGCG GTTCCCGCG
201 CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
251 TGCCGGTTCG CTGGCTGTAT GGTGCGCCGT CTTATCAGAT AGTCGGTTCG
301 ATATTGGAAA GCAATCCTGC CGAGGCGCGT GAATTGTGCG GCAATCTTCC
351 CGGTTCGCTT TATTTGTGCG AGGCATTATT TTTTATTTT GGCTTGACAG
401 TTTGGAGATA TTGTGTATCG GGGGGGGTAT TTGCTGACGT AAAAACTAT
451 AAACGCCGCA GCAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTC
501 CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTG CGAGAACCTG
551 ATGCCGGCCT GTTGTGAAT ATTTTCGACC TGTATTACGA TTTGGCTTCC
601 .GCGCCGCA CAATATGCCG CCAAGCGCG CCACATTTTG GAAGCAGCAA
651 AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA
  
```

This corresponds to the amino acid sequence <SEQ ID 264; ORF 084.a>:

a084.pep

```

1  MKQSARIKMN DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
51  ALASVVFLLL LARGFPRVSS VLLLIYVGT ALYLPVGWLY GAPS YQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWRYCVS GGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDA GLLN IFDLYYDLAS
201 XAGTICRQAR PHFGSSKKSVMAYPSCCAQ V*
  
```

m084/a084 92.2% identity over a 231 aa overlap

	10	20	30	40	50	60
m084.pep	MKQSARIKXMQTLTYTLGICALLTFXXXXXXHYHPEY EYGYRYSAVGALASVVFLLL					
a084	MKQSARIKNMDQTLKNTLGIALLAFCFGA AIASGYHLEY EYGYRYSAVGALASVVFLLL					
	10	20	30	40	50	60
m084.pep	LARGFPRVSSVLLIYVGTALYLPVGWLYGAPS YQIVGS ILESNPAEAREFVGNLPGSL					
a084	LARGFPRVSSVLLIYVGTALYLPVGWLYGAPS YQIVGS ILESNPAEAREFVGNLPGSL					
	70	80	90	100	110	120
m084.pep	YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL					
a084	YFVQALFFIFGLTVWRYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL					
	130	140	150	160	170	180
m084.pep	REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVMAYPSCCAQVX					
a084	REPDAGLLLNIFDLYYDLASXAGTICRQARPHFGSSKKSVMAYPSCCAQVX					
	190	200	210	220	230	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 265>:

g085.seq

```

1  ATGGGCAAAG GGCAGGACTT CACGCCCTG CGCGACGCGT TGAAAGATAA
  
```

```

51  GGCAAAAGGC GTGTTCTGA TCGGCGTCGA TGCGCCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AACCTGACCG ACTGCGTCAC TTTGGAAGAG
151 GCGGTTTCTG CCGCATACGC CCAAGCCGAA GCGGGCGATA TTGTCTTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
251 CGGAAGTGTT tatCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 266; ORF 085.ng>:

**g085.pep**

```

1  MGKGQDFTPL RDALKDKAKG VFLIGVDAPO IRRDLGCGCL NLTDCVTLEE
51  AVQTAYAQA EAGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 267>:

**m085.seq**

```

1  ATGGGTAAAG GGCAGGACTT CACGCCCCTG CGCGATGCAC TGGTAGGCAA
51  GGCAAAAGGC GTGTTCTTGA TTGGTGTCGA TGCGCCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AATATGACCG ACTGCGCCAC TTTGGGAGAA
151 GCCGTTTCTG CCGCATATGC CCAAGCCGAA GCAGGCGATA TTGTGTTGCT
201 CAGCCCCGCC TCGCGAGCTT TTGATATGTT CAAAGGCTAC GCGCACCGTT
251 CGGAAGTGTT TATCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 268; ORF 085>:

**m085.pep**

```

1  MGKGQDFTPL RDALVGKAKG VFLIGVDAPO IRRDLGCGCL NMTDCATLGE
51  AVQTAYAQA EAGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 085 shows 94.7% identity over a 94 aa overlap with a predicted ORF (ORF 085.ng) from *N. gonorrhoeae*:

**m085/g085**

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGCLNMTDCATLGEAVQTAYAQA					
g085	MGKGQDFTPLRDALKDKAKGVFLIGVDAPOIRRDLGCGCLNLTDCVTLEEAVQTAYAQA					
	10	20	30	40	50	60
	70	80	90			
m085.pep	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
g085	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
	70	80	90			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 269>:

**a085.seq**

```

1  ATGGGCAAAG GGCAGGACTT CACGCCCCTG CGCGACGCGC TTGCCGGCAA
51  GGCAAAAGGC GTGTTCTTGA TCGGTGTCGA TGCGCCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGATCTG AATATGACCG ACTGCGCCAC TTTGGAAGAA
151 GCGGTTTCTG AGGCATATGC CCAAGCCGAA GCGGGCGATA TCGTGCTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
251 CGGAAGTGTT TATCGGGGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 270; ORF 085.a>:

**a085.pep**

```

1  MGKGQDFTPL RDALAGKAKG VFLIGVDAPO IRRDLGCDL NMTDCATLEE
51  AVQKAYAQA EAGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

**m085/a085** 94.7% identity over a 94 aa overlap

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGCLNMTDCATLGEAVQTAYAQA					

```

a085      |||||:|||||
MGKGQDFTPLRDALAGKAKGVFLIGVDAPQIRRLDGLGCDLNMTDCATLEEAVQKAYAQAEE
          10      20      30      40      50      60

          70      80      90
m085.pep  AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
          |||||
a085      AGDIVLLSPACASFDMFKGYAHRSEVFIGAFKALX
          70      80      90

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 271>:

g086.seq

```
1 ATGGTGGTGC TGATGACGGC GTTCGGCCTG CTGATGATT ATTTCGGCTTC
51 TGTGATTTTG GCATCGAAGG AAGGCGGCGA TCAGTTTTTTC TATTTGACCA
101 GGCAGGCGGG GTTCGTCGTT GCCCGCCTTA TAGCGAGCGG TTTTTTATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTGTGTGCCGT GGATTTTTTGC
201 CTTATCCGGC CTGTTGTCTG TAGCCGTATT GATTGCCGGG CCGGAAATCA
251 ATGGCGCGAC CCGTTGGATA CTTTGGGTC CGTTGAATT CCAGCCGACC
301 GAGCTGTTCA AGCTGGCAGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
351 CCGTGAAGAA GTGTTGCGCA GCATGAAAG TTTGGGTTGG CAGTCGATTT
401 GGCGGGGGAC GGCCAACCTG ATTATGTCCG CCACCAATCC GCAGGCACGT
451 CGTGAACAT TAGAAATGTA CGgCGTTTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGT GCGTTCGGTT TGGTGCTAG AATGGTACAG CCGGATTTTCG
551 GTTCGTTTGT CGTCATTACC GTCATTACCG TTGGAATGCT GTTCTGGCA
601 GGATTGCCGT GGAAATATTT TTTTGTCTCG GTAGGCAGCG TCTTGGGTGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTTGGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
801 TGCGAGTTTG AGCAAACGCG GCTTCTGCC GGAAGCGCAT ACCGATTTTA
851 TTTTGGCCAT CATCGCTGAA GAATTCGGCT TCTTCGGGAT GTGCGTGCTG
901 ATATTCTGTT ACGGCTGGCT GGTGGTGC GGCTTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGtttgA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTTGGAT CCGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCGGA CCAAAGSTCT GACGctGcG tTGATGTCTT ATGGcggTTC
1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTTG CGTATCGATT
1151 ATGAAAACCG CCAGAAAATG CGCGGTTACC GGGTGGAGTA AA
```

This corresponds to the amino acid sequence <SEQ ID 272; ORF 086.ng>:

g086.pep

1	MVVLMTAFGL	LMIYSASVYL	ASKEGGDQFF	YLTRQAGFVV	AGLIASGFLW
51	<u>FLCRMRTWRR</u>	<u>LVPWIFALSG</u>	<u>LLLVAVLIAG</u>	REINGATRWI	PLGPLNFQPT
101	ELFKLAVILY	LASLFTREEE	VLRSMESLGW	QSIWRGTANL	IMSATNPQAR
151	RETLEMYGRF	RAIILPIMLV	AFGLVLIMVQ	PDFGSFVVIT	VITVGMLFLA
201	<u>GLPWKYFFVL</u>	<u>VGSVLGGMVL</u>	<u>MITAAPYRVQ</u>	RVVAFDPWK	DPQGAGYQLT
251	HSLMAIGRGE	WFGMGLGASL	SKRGFLPEAH	TDIFAIIAIE	EFGFGGMCVL
301	IFCYGWLVVR	AFSIGKQSRD	LGLTFNAYIA	SGIGIWIGIQ	SFNIGVNI
351	<u>ALPTKGLTLP</u>	<u>LMSYGGSSVF</u>	<u>FMLISMMLLL</u>	RIDYENRQKM	RGYRVE*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 273>>:

m086.seq

```
1 ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
51 TGTGTATTTG GCATCAAAGC AAGGCGGCGA TCAGTTTTTC TATTTGACCA
101 GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCCT GGATTTTTTG
201 CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG GCGCAAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
301 GAGCTGTTCA AGCTGGCGGT CATCCTTTAT TTGGCAAGCC TGTTACGCG
351 CCGTGAAGAA GTGTTGcGCA GCATGGAAGC TTTGGGTTGG CAGTCGATTT
401 GCGGGGGGAC GGCCAATCTG ATCATGTC CGACCAATCC GCAGrCACGT
451 CGTGAAaACAT TAGAAATGTA CGGCCGTwTC CCGCGGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCC
```



```

551 GTTCGTTTGT CGTCATTACC GTCATTGCCG TTGGAATGCT GTTTTTGGCA
601 GGATTGCCGT GGAAATATTT TTTCGTCCTG GTAGGCAGCG TCTTGGGCGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTTTGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
801 TGCAGAGTTG AGCAAACGCG GCTTCTGCC GGAAGCGCAT ACCGATTTTA
851 TTTTGGCCAT CATCGCCGAA GAATTCGGTT TCTTCGGTAT GTGCGTGCTG
901 ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGTTTGA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTTGGAT CGGkrTCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCCGA mCAAAGyCT GACGyGCCG Tg .AtGTCCw ATGGCGGTTT
1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGtKg CGTATAGATT
1151 ATGAAAACCG CCGGAAAATG CGCGGTTATC GGGTGGAGTA A

```

This corresponds to the amino acid sequence <SEQ ID 274; ORF 086>:

```

m086.pep
  1  MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW
 51  FLCRMRTWRR LVPWIFALSG LLLVVVLIAG REINGATRWI PLGPLNFQPT
101  ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQXR
151  RETLEMYGRX RAILPIMLV AFGLVLIMVQ PDFGSFVVIT VIAVGMFLFA
201  GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFLDPWK DPQAGYQLT
251  HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIE EFGFFGMCVL
301  IFCYGWLVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGXQ SFFNIGVNIG
351  ALPXXGLTXP XMSXGGSSVF FMLISMMLLX RIDYENRRKM RGYRVE*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 086 shows 96.7% identity over a 396 aa overlap with a predicted ORF (ORF 086.ng) from *N. gonorrhoeae*:

m086/g086

	10	20	30	40	50	60
m086.pep	MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW FLCRMRTWRR					
g086	MVVLMTAFGL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLW FLCRMRTWRR					
	10	20	30	40	50	60
m086.pep	LVPWIFALSGL LLLVVVLIAG REINGATRWI PLGPLNFQPT ELFKLAVILY LASLFTRREE					
g086	LVPWIFALSGL LLLVAVLIAG REINGATRWI PLGPLNFQPT ELFKLAVILY LASLFTRREE					
	70	80	90	100	110	120
m086.pep	LVPWIFALSGL LLLVVVLIAG REINGATRWI PLGPLNFQPT ELFKLAVILY LASLFTRREE					
g086	LVPWIFALSGL LLLVAVLIAG REINGATRWI PLGPLNFQPT ELFKLAVILY LASLFTRREE					
	70	80	90	100	110	120
m086.pep	VLRSMESLGW QSIWRGTANL IMSATNPQXR RETLEMYGRX RAILPIMLV AFGLVLIMVQ					
g086	VLRSMESLGW QSIWRGTANL IMSATNPQAR RETLEMYGRF RAILPIMLV AFGLVLIMVQ					
	130	140	150	160	170	180
m086.pep	VLRSMESLGW QSIWRGTANL IMSATNPQXR RETLEMYGRX RAILPIMLV AFGLVLIMVQ					
g086	VLRSMESLGW QSIWRGTANL IMSATNPQAR RETLEMYGRF RAILPIMLV AFGLVLIMVQ					
	130	140	150	160	170	180
m086.pep	PDFGSFVVIT VIAVGMFLFAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
g086	PDFGSFVVIT VITVGMLFAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
	190	200	210	220	230	240
m086.pep	PDFGSFVVIT VIAVGMFLFAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
g086	PDFGSFVVIT VITVGMLFAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
	190	200	210	220	230	240
m086.pep	DPQAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
g086	DPQAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
	250	260	270	280	290	300
m086.pep	DPQAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
g086	DPQAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
	250	260	270	280	290	300
	310	320	330	340	350	360

```

m086.pep      IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGXQSFFNIGVNIGALPXXKGLTXP
                |||
g086           IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGIQSFFNIGVNIGALPTKGLTLP
                |||
                310      320      330      340      350      360

                370      380      390
m086.pep      XMSXGGSSVFFMLISMMLLXRIDYENRRKMRGYRVEX
                |||
g086           LMSYGGSSVFFMLISMMLLLRIDYENRQKMRGYRVEX
                |||
                370      380      390

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 275>:

## a086.seq

```

1   ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
51  TGTGTATTTC GCATCAAAAG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
101 GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
151 TTTCTTTGCA GGTAGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTCG
201 CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
301 GAGCTGTTCA AGCTGGCGGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
351 CCGTGAAGAA GTGTTGCGCA GCATGGAAG TTTGGGTTGG CAGTCGATTT
401 GGCGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGGCACGT
451 CGTGAACAT TAGAAATGTA CGGCCGTTTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTCTGAT AATGGTACAG CCGGATTTTCG
551 GTTCGTTTGT CGTCATTACC GTCATTGCCG TTGGAATGCT GTTTTGGGCA
601 GGATTGCCGT GGAAATATTT TTTGTCCTG GTAGGCAGCG TCTTGGGCGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTTGGA CCCGTGGAAG GACCCGAGG GTGCCGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGTA TGGGTTTGGG
801 TGCAGAGTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
851 TTTTGGCCAT CATCGCCGAA GAATTCGGTT TCTTCGGTAT GTGCGTGTCTG
901 ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGTTTGA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTTGGAT CCGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCCGA CCAAAGTCT GACGCTGCCG TTGATGTCCT ATGGCGGTTT
1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTTG CGTATAGATT
1151 ATGAAAACCG CCGGAAAATG CGCGTTTACC GGGTGGAGTA A

```

This corresponds to the amino acid sequence <SEQ ID 276; ORF 086.a>:

## a086.pep

```

1   MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW
51  FLCRMRTWRR LVPWIFALSG LLLVVVLIAG REINGATRWI PLGPLNFQPT
101 ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQAR
151 RETLEMYGRF RAAILPIMLV AFGLVLIMVQ PDFGSFVVIT VIAVGMLFLA
201 GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFDPWK DPQAGYQLT
251 HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIE EFGFFGMCVL
301 IFCYGWLVV R AFSIGKQSRD LGLTFNAYIA SGIGIWIGIQ SFFNIGVNIG
351 ALPTKGLTLP LMSYGGSSVF FMLISMMLLL RIDYENRRKM RGYRVE*

```

m086/a086 98.0% identity over a 396 aa overlap

```

                10      20      30      40      50      60
m086.pep      MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
                |||
a086           MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
                |||
                10      20      30      40      50      60

                70      80      90      100     110     120
m086.pep      LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTLFLKLAIVLYLASLFTRREE
                |||
a086           LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTLFLKLAIVLYLASLFTRREE
                |||

```

277

	70	80	90	100	110	120
m086.pep	130	140	150	160	170	180
	VLRSMESLGWQSIWRGTANLIMSATNPQXRRETLEMYGRXRAIILPIMLVAFGLVLIMVQ					
a086	VLRSMESLGWQSIWRGTANLIMSATNPQARRETLEMYGRFRAIILPIMLVAFGLVLIMVQ					
	130	140	150	160	170	180
m086.pep	190	200	210	220	230	240
	PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
a086	PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
	190	200	210	220	230	240
m086.pep	250	260	270	280	290	300
	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
a086	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL					
	250	260	270	280	290	300
m086.pep	310	320	330	340	350	360
	IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGXQSFFNIGVNIGALPXXGLTXP					
a086	IFCYGWLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGIQSFFNIGVNIGALPTKGLTLP					
	310	320	330	340	350	360
m086.pep	370	380	390			
	XMSXGGSSVFFMLISMMLLXRIDYENRRKMGRGYRVEX					
a086	LMSYGGSSVFFMLISMMLLLRIDYENRRKMGRGYRVEX					
	370	380	390			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 277>:

g087.seq

1	ATGGGCGGTA	AAACCTTTAT	GCTGATGGCG	GGCGGAACGG	GCGGACACAT
51	TTTCCCAGCT	CTGGCTGTGG	CGGATTCATT	GCGCGTGC	GGTCATCATG
101	TAATTTGGCT	GGGCAGCAAG	GATTTCGATG	AAGAGCGCAT	CGTGCCGCAA
151	TACGGCATA	GCTTGGAAC	GCTGCGGATT	AAAGGAATAC	GCGGCAACGG
201	CATCAAACGC	AAGCTGATGC	TTCCGTTTAC	TCTGTACAAA	ACCGTCCGCG
251	AAGCGCAGCG	GATTATCCGC	AAACACCGTG	TCGAGTGC	CATCGGCTTC
301	GGCGGTTTTG	TTACCTTTCC	CGGCGGTCTG	GCGGCGAAAC	TCTTGGGCGT
351	GCCGATTGTG	ATTCACGAGC	AAAACGCCGT	GGCAGGCTTG	TCCAACCGCC
401	AccTGTTCG	ctGGGCGAAA	CGGGTGTGT	ACGCTTTTCC	GAAAGCGTTC
451	AGCCACGAAG	GCGGTTTGGT	CGGCAACCCC	GTCCGCGCCG	ATATTAGCAA
501	CTGCCCCGTG	CCTGCCGAAC	GCTTCCAAGG	GCGCGAAGGC	CGTCTCAAAA
551	TTTTGGTGGT	CGGCGGCAGT	TTGGGTGCGG	ACGTTTGTAA	CAAAACCGTA
601	CCGCAGGCGT	TGGCACTGCT	GCCTGAAGAG	GTGCGCCCGC	AGATGTACCA
651	CCAGTCGGGG	CGTAACAAGC	TGGGCAATCT	TCAGGCGGAT	TATGACGCGT
701	TGGGCGTGAA	AGCGGAATGC	GTGGAATTTA	TTACCGACAT	GGTGTCCGCC
751	TACCGTGATG	CCGATTGGT	GATTGTCCGT	GCCGGCGCGC	TGACGATTGC
801	CGAGTTGACG	GCGGCGGGC	TGGGCGCGTT	GTTAGTGCCG	TATCTCAGC
851	CCGTTGATGA	CCATCAAACC	GCCAACGCGC	GTTTCATGGT	GCAGGCAGAA
901	GCGGGGCTGC	TGTTGCCGCA	AACCCAGTTG	ACGGCGGAAA	AACTCGCCGA
951	AATCCTCGGC	AGCCTCAACC	GCGAAAAATG	CCTCAAATGG	GCGGAAAACG
1001	CCCGTACGTT	GGCATTGCCG	CACAGCGCGG	ATGACGTTGC	CGAAGCCGCG
1051	ATTGCGTGTG	CGGCGTAA			

This corresponds to the amino acid sequence <SEQ ID 278; ORF 087.ng>:

g087.pep

1	MGGKTFMLMA	GGTGGHIFPA	LAVADSLRVR	GHHVIWLGSK	DSMEERIVPQ
51	YGIRLET Lai	KGIRNGIKR	KLMLPFTLYK	TVREAQRIIR	KHRVECVIGF
101	GGFVTFPGGL	AAKLLGVPIV	IHEQNAVAGL	SNRHLSRWAK	RVLYAFP KAF
151	SHEGGLVGNP	VRADISNLPV	PAERFQREG	RLKILVVGGS	LGADV LNKTV

201 PQALALLPEE VRPQMYHQSG RNKLGNIQAD YDALGVKAEC VEFITDMVSA  
 251 YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHOT ANARFMVQAE  
 301 AGLLLPQTQL TAEKLAELILG SLNREKCLKW AENARTLALP HSADDVAEAA  
 351 IACAA\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 279>:

m087.seq  
 1 ATGGGCGGTA AAACCTTTAT GCTGAwkkCG GCGGGAACGG GCGGACATAT  
 51 TTTCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG  
 101 TGATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGTAT CGTGCCGCAA  
 151 TACGGCATAC GCTTGGAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG  
 201 CATCAAACGC AAACGTATGC TGCCGGTTAC TTTGTATCAA ACCGTCCGCG  
 251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCCT CATCGGCTTC  
 301 GCGGCTTCG TTACCTTCCC CGGCGGTTTG GCGGCGAAGC TATTArGCGT  
 351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC  
 401 ACCTGTTCGCG CTGGGCGAAG CGGGTGTGTG ACGCTTTTCC GAAAGCGTTC  
 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA  
 501 CCTGCCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA  
 551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTGTAA CAAAACCGTA  
 601 CCGCATGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC ATATGTACCA  
 651 CCAATCGGGA CGGGGCAAGC TGGGCATCTT GCAGGCGnnn nnnnnnnnnn  
 701 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  
 751 nnnGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG CGGTTGACGA  
 801 TCACCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG GCGGGATTGC  
 851 TGTTCGCCGA AACCCAGTTG ACGGCGGAAA AACTCGCCGA GATTCTCGGC  
 901 GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG CCCGTACGTT  
 951 GGCACGTCCG CACAGTGCGG ACGACGTGGC GGAAGCCGCG ATTGCGTGTG  
 1001 CGGCGTAA

This corresponds to the amino acid sequence <SEQ ID 280; ORF 087>:

m087.pep  
 1 MGGKTFMLXX GGTGGHIFPA LAVADSLRAR GHVHVIWLGSK DSMEERIVPQ  
 51 YGIRLETIAI KGVVRNGIKR KLMLPVTLYQ TVREAQRIIR KHRVECVIGF  
 101 GGFVTFPGGL AAKLLXVPIV IHEQNAVAGL SNRHLSRWAK RVLVYAFPKAF  
 151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGS LGADVLNKTIV  
 201 PHALALLPDN ARPHMYHQSG RGKLGILQAX XXXXXXXXXXX XXXXXXXXXXX  
 251 XAGLGALLVP YPHAVDDHOT ANARFMVQAE AGLLLPQTQL TAEKLAELILG  
 301 GLNREKCLKW AENARTLALP HSADDVAEAA IACAA\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 087 shows 83.9% identity over a 355 aa overlap with a predicted ORF (ORF 087.ng) from *N. gonorrhoeae*:

m087/g087

	10	20	30	40	50	60
m087.pep	MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVHVIWLGSKDSMEERIVPQYIGIRLETIAI					
g087	MGGKTFMLMAGGTGGHIFPALAVADSLRVRGHHVHVIWLGSKDSMEERIVPQYIGIRLETIAI					
	10	20	30	40	50	60
m087.pep	KGVVRNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVPIV					
g087	KGIRNGIKRKLMLPFTLYKTVEAQRIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV					
	70	80	90	100	110	120
m087.pep	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
g087	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
	130	140	150	160	170	180
m087.pep	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
g087	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
	130	140	150	160	170	180

279

	190	200	210	220	229
m087.pep	RLKILVVGGS	LGADVLNKT	VPHALALLP	DNARPHMY	HQSGRGKLGILQA-----
g087	RLKILVVGGS	LGADVLNKT	VPQALALLP	EEVRPQMY	HQSGRNKLG
	190	200	210	220	230
					240
					250
m087.pep	-----	AGLGALLV	PYPH	AVDDHQT	ANARFMVQAE
g087	VEFITDMV	SAYRDADL	VICRAGAL	TIAELTA	AGLGALLV
	250	260	270	280	290
					300
					310
m087.pep	AGLLLPQT	QLTAEKLA	EILGGLN	REKCLKW	AEENARTL
g087	AGLLLPQT	QLTAEKLA	EILGSLN	REKCLKW	AEENARTL
	310	320	330	340	350

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 281>:

a087.seq

1	ATGGGCGGTA	AAACCTTTAT	GCTGATGGCG	GGCGGAACGG	GCGGACATAT
51	TTTCCCCGCG	CTGGCGGTGG	CGGATTCATT	GCGCGCGCGC	GGCCATCATG
101	TAATTTGGCT	GGGCAGCAAG	GATTCGATGG	AAGAGCGCAT	CGTGCCGCAA
151	TACGACATCC	TGCTCGAAAC	GCTGGCGATT	AAAGGCGTGC	GCGGCAACGG
201	CATCAAACGC	AAGCTGATGC	TGCCGTTTAC	TTTGTATCAA	ACTGTCCGCG
251	AAGCGCAGCA	GATTATCCGC	AAACACCGTG	TCGAGTGCGT	CATCGGCTTC
301	GGCGGCTTCG	TTACCTTTCC	CGGCGGTTTG	GCGGCGAAGT	TATTAGGCGT
351	GCCGATTGTG	ATTCACGAGC	AAAACGCCGT	GGCAGGTTTG	TCCAACCGCC
401	ACCTGTGCGC	CTGGGCGAAG	CGGGTGTGTG	ACGCTTTTCC	GAAAGCGTTC
451	AGCCACGAAG	GCGGCTTGGT	CGGCAACCCC	GTCCGCGCCG	ATATTAGCAA
501	CCTGCCCCGT	CCTGCCGAAC	GCTTCCAAGG	GCGTGAAGGC	CGTCTGAAAA
551	TTTTGGTGGT	CGGCGGCACT	TTGGGCGCGG	ACGTTTTGAA	CAAAACCGTA
601	CCGCAGGCAT	TGGCTTTGCT	GCCCGACAAT	GCGCGTCCGC	AGATGTACCA
651	CCAATCGGGA	CGGGGCAAGC	TGGGCAGCTT	GCAGGCGGAT	TACGACGCGC
701	TGGGCGTGCA	AGCGGAATGC	GTGGAATTTA	TTACCGATAT	GGTGTCCGCC
751	TACCGCGATG	CCGATTTGGT	GATTTGCCGT	GCCGCGCGCG	TGACGATTGC
801	CGAGTTGACG	GCGGCGGGAT	TGGGTGCGTT	GTTAGTGCCG	TATCCTCACG
851	CCGTTGATGA	CCATCAAACC	GCCAACGCGC	GTTTTATGGT	GCAGGCGGAG
901	GCGGGATTGC	TGTTGCCGCA	AACCCAGTTG	ACGGCGGAAA	AACTCGCCGA
951	GATTCTCGGC	GGCTTAAACC	GCGAAAAATG	CCTCAAATGG	GCAGAAAACG
1001	CCCGTACGTT	GGCACTGCCG	CACAGTGCGG	ACGACGTTGC	CGAAGCCGCG
1051	ATTGCGTGTG	CGGCGTAA			

This corresponds to the amino acid sequence <SEQ ID 282; ORF 087.a>:

a087.pep

1	MGGKTFMLMA	GGTGGHIFPA	LAVADSLRAR	GHHVIWLGSK	DSMEERIVPQ
51	YDILLETLAI	KGVRNGNIKR	KLMLPFTLYQ	TVREAQQIIR	KHRVECVIGF
101	GGFVTFPGL	AAKLLGVPIV	IHEQNAVAGL	SNRHLSRWAK	RVLYAFPKAF
151	SHEGGLVGNP	VRADISNLPV	PAERFQREG	RLKILVVGGS	LGADVLNKT
201	PQALALLPDN	ARPQMYHQS	RGKLGSLQAD	YDALGVQAE	VEFITDMVSA
251	YRDADLVICR	AGALTIAELT	AAGLGALLVP	YPHAVDDHQT	ANARFMVQAE
301	AGLLLPQTQL	TAEKLAELG	GLNREKCLKW	AENARTLALP	HSADDVAEAA
351	IACAA*				

m087/a087 85.4% identity over a 355 aa overlap

	10	20	30	40	50	60
m087.pep	MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQY	GIRLETLAI				
a087	MGGKTFMLMAGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQY	DILLETLAI				
	10	20	30	40	50	60

m087.pep	70	80	90	100	110	120
	KGVRGNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVPIV					
a087	KGVRGNGIKRKLMLPFTLYQTVREAQQIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV					
	70	80	90	100	110	120
m087.pep	130	140	150	160	170	180
	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
a087	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
	130	140	150	160	170	180
m087.pep	190	200	210	220	230	240
	RLKILVVGGS LGADV LNKT VPHALALLPDNARPHMYHQS GRGKLGILQAXXXXXXXXXXX					
a087	RLKILVVGGS LGADV LNKT VQALALLPDNARPMYHQS GRGKLGSLQADYDALGVQAE					
	190	200	210	220	230	240
m087.pep	250	260	270	280		
	XX-----XXXXXXXXXAGLGALLVPYPHAVDDHQTANARFMVQAE					
a087	VEFITDMVSAYRDADLVICRAGALTIAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE					
	250	260	270	280	290	300
m087.pep	290	300	310	320	330	
	AGLLLPQTQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					
a087	AGLLLPQTQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 283>:

g088.seq

```

1  ATGTTTTTAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
51  TTTTCAATAC ACCACATTCC GCGCCGTTAT GGCGGCGTTG ACCGCCTTGG
101 CGTTTTCCCT GATGTTCCGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
151 AAATGCGGGC AGGCAGTGCG TACCGACGGC CCGCAAACCC ACCTCGTCAA
201 AAACGGCAGC CCGACGATGG GCGGTTTCGCT GATTCTGACC GCCATTACCG
251 TGTCCACCTT GTTGTGGGGC AACTGGGCGA ACCCGTATAT CTGGATTCTC
301 TTGGGCGTAC TGCTTGCCAC CCGTGCGCTC GGTTTTTACG ACGACTGGCG
351 CAAAGTCGTT TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAATGG
401 TGTGGCAGTC AAGCGTTGCC GTTatcgCG GTttggcaTT GTTTTACctt
451 gCcgCAATT CCGCAACAA TATTTTGATT GTCCCGtttT TCAAACAAAT
501 CGCCCTGCCG CTGGGCGTGG TCGGCTTttt gGtggtgTCT TACCTGACCA
551 TCGTCGGCAC ATCCAACGCC GTCAACCTCA CcgaCGGCTT GGACGGCCTT
601 GCCGCcttcc cgttcgctcc cgttgccgc GGGCTCGCCA ttttcgcctA
651 CGTCAGCGGA CACTACCAAT TTTCCCAATA CCTCCAGCTT CCCTATGTCTG
701 CCGGCGCGAA CGAAGTCGCT ATATTCTGCA CCGCATATGT CCGCGCGTGC
751 CTCGGATTTT TGTGGTTCAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTCGGCGCG CTGGCATTGG GTGCCGCGCT CGGTaccGtt gCCGTcaTcg
851 tCCGCCAAGA ATTTGTcctc gtcattatGG GCGGTCTGTT cgtcgtagaa
901 gccgtgTCCG TTATGCTTCa tgtcggCTGG TACAAGAAAA Ccaaaaaacg
951 CATCTTcTcTg acgGcaccga ttcattacca ttaCCaactt cgatgCTGGa
1001 aagaaacgca agtcgtcgtc CGTTtCTGGA TTAtTAccat cgtcgtggtt
1051 tTgataggtt tGagtacctt caAAattcgc ggaaactatg ccgTCCGAAC
1101 ACCTTTTACA CGGCATTTGA ACGCGCAATA A

```

This corresponds to the amino acid sequence <SEQ ID 284; ORF 088.ng>:

g088.pep

```

1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
51  KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL
101 LGVLLATGAL GFYDDWRKVY YKDPNGVSAK FKMVWQSSVA VIAGLALFYL
151 AANSANNILI VPFFKQIALP LGVVGFVLVS YLTIVGTSNA VNLTGDLGL

```

201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC  
251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE  
301 AVSVMLHVGW YKTKTKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV  
351 LIGLSTLKIR GNYAVRTPFR RHLNAQ\*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 285>:

```
m088.seq
1 ATGTTTTTAT GGCTCGCACA TTTCAGCAnC TGGTTAACCG GTCTGAATnn
51 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
101 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
151 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
201 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
251 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
301 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
351 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
401 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
451 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
501 nnnnnnnnnnnn nnnGGCGTGG TCGGCTTTTT GGTGTTGTCT TACCTGACCA
551 TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
601 GCGACCTTCC CCGTCGTCCT CGTTGCCGCC GGCCTCGCCA TCTTCGCCTA
651 TGCCAGCGGC CACTCACAAT TTGCCCAATA CTTGCAATTA CCTTACGTTG
701 CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CCGCGCGTGC
751 CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTCCGTGCA TTGGCATTGG GTGCCGCGCT CGGTACCGTC GCGGTTATCG
851 TCCGCCAAGA GTTTGTCCCTC GTCATTATGG GCGGATTATT TGTCGTAGAA
901 GCCGTATCCG TTATGCTTCA GGTGGCTGG TATAAGAAAA CCAAAAAACG
951 CATCTTCTTG ATGGCGCCCA TCCATCACCA CTACGAACAA AAAGGCTGGA
1001 AAGAAACCCA AGTCGTCGTC CAGCTTTTGA TTATTACCAT CGTCTTGGTG
1051 TTGATCGGTT TGAGTACCTT CAAAATCCGC TGAACCTATG CCGTCTGAAC
1101 ATCTTTTACA CGGCATTTGA ACGCGCAATA A
```

1	MFLWLAHFSN	WLTGLNIFOY	TTFRAVMAAL	TALAFSLMFG	PWTIRRLTAL
51	KCGQAVRTDG	PQTHLVKNGT	<u>PTMGGSLILT</u>	AITVSTLLWG	NWANPYIWIL
101	<u>LGVLATGAL</u>	GFYDDWRKVV	YKDPNGVSAK	FKMVWQSSVA	VIAGLALFYL
151	<u>AANSANNILI</u>	VPFFKQIALP	<u>LGVVGFVLVS</u>	YLTIVGTSNA	VNLTGDGLDGL
201	<u>AAFPFVLVAA</u>	GLAIFAYVSG	HYQFSQYLQL	PYVAGANEVA	IFCTAMCGAC
251	<u>LGFLVFNAYP</u>	AQVFMGDVGA	<u>LALGAALGTV</u>	AVIVRQEFVL	VIMGGFLVVE
301	<u>AVSVMLHVGW</u>	YKKTGKRIFL	TAPIHHHYQL	RCWKETQVVV	RFWIITIVVV
351	LIGLSTLKIR	GNAYVRTPPR	RHLNAQ*		

This corresponds to the amino acid sequence <SEQ ID 286; ORF 088>:

```
m088.pep
1  MFLWLAHFSX WLTGLNXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
51  XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
101 XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
151 XXXXXXXXXXXX XXXXXXXXXXXX XGVVGFVLVLS YLTIVGTSNA VNLTDLGLDGL
201 ATFPVVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEV IFCTAMCGAC
251 LGFLWFNAYP AQVFMGDVGA LALGAALGT AVIVRQEFVL VIMGGLFVVE
301 AVSVMLQVGW YKTKKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
351 LIGLSTLKIR XTYAVXTSFR RHLNAQ*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 088 shows 91.7% identity over a 205 aa overlap with a predicted ORF (ORF 088.ng) from *N. gonorrhoeae*:

m088/g088

m088.pep  
 GVVGFLVLSYLTIVGTSNAVNLTGDLGLA  
 |||||  
 g088  
 IAGLALFYLAANSANNILIVPFFKQIALPLGVVGFLVLSYLTIVGTSNAVNLTGDLGLA  
 150 160 170 180 190 200

	40	50	60	70	80	90
m088.pep	TFPVVLVAAGLAI FAYASGHSQFAQYLQLPYVAGANEVVFICTAMCGACLGFLWFNAYPA					
	:                    :                    :					
g088	AFPFVLVAAGLAI FAYVSGHYQFSQYLQLPYVAGANEVAIFCTAMCGACLGFLWFNAYPA					
	210	220	230	240	250	260

	100	110	120	130	140	150
m088.pep	QVFMGDVGALALGAALGTVAIVIRQEFVLVIMGGLFVVEAVSVMLQVGWYKTKKRIFLM					
g088	QVFMGDVGALALGAALGTVAIVIRQEFVLVIMGGLFVVEAVSVMLHVGWYKTKKRIFLT					
	270	280	290	300	310	320

	160	170	180	190	200
m088.pep	APIHHHYEQKGWKETQVVVRFWIIITIVLVLIGLSTLKIRXTYAVXTSFRRLHNAQX				
	: :				
g088	APIHHHYQLRCWKETQVVVRFWIIITIVVLIGLSTLKIRGNVAVRTPFRRLHNAQX				
	330	340	350	360	370

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 287>:

a088.seq

```

1  ATGTTTTTAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
51  TTTTCAATAC ACCACATTCC GCGCCGTCAT GCGCGCGTTG ACCGCGTTGG
101 CGTTTCCCTT GATGTTTCGGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
151 AAATGCGGGC AGGCAGTGCG TACCGACGGT CCGCAAACCC ACCTCGTCAA
201 AAACGGCACG CCGACGATGG GCGGTTGCTG GATTCTGACC GCCATTACCG
251 TGTCACCCCT GTTGTGGGGC AACTGGGCAA ACCCGTATAT CTGGATTCTC
301 TTGGGCGTAT TGCTCGCCAC GGGCGCACTC GGTTTTACG ACGACTGGCG
351 CAAAGTCGTC TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAAATGG
401 TGTGGCAGTC AAGCGTTGCC ATTATCGCCG GTTTGGCATT GTTTTACCTT
451 GCCGCCAATT CCGCCAACAA TATTTTGATT GTCCCGTTCT TCAAACAAAT
501 CGCCCTGCCG CTGGGCGTGG TCGGCTTTTT GTGTTGTCT TACCTGACCA
551 TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
601 GCGACCTTCC CCGTCGTCTT CGTTGCCGCC GGCTCGCCA TCTTCGCCTA
651 TGCCAGCGGC CACTCACAAT TTGCCCAATA CCTGCAATTA CTTACGTTG
701 CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CGGCGCGTGC
751 CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTGCGTGCA TTGGCATTGG GTGCCGCGCT CGGTACCGTC GCCGTCATCG
851 TCCGCCAAGA GTTTGTCCTC GTCATTATGG GCGGATTATT TGTCTAGAA
901 GCCGTATCCG TTATGCTTCA GGTGCGCTGG TATAAGAAAA CCAAAAAACG
951 CATCTTCCTG ATGGCGCCCA TCCATCACCA CTACGAACAA AAAGGCTGGA
1001 AAGAAACCCA AGTCGTCGTC CGCTTTTGGG TTATTACCAT CGTCTTGGTG
1051 TTGATCGGTT TGAGTACCCT CAAAATCCGC TGAACCTATG CCGTCTGAAC
1101 ACCTTTCAGA CGGCATTGTA ACGCGCAATA A

```

This corresponds to the amino acid sequence <SEQ ID 288; ORF 088.a>:

a088.pep

```

1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
51  KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL
101 LGVLLATGAL GFYDDWRKVY YKDPNGVSAK FKMVWQSSVA IIAGLALFYI
151 AANSANNILI VPFFKQIALP LGVVGFVLVS YLTIVGTSNA VNLTDGLDGL
201 ATFPVVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEVV IFCTAMCGAC
251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
301 AVSVMLQVGW YKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
351 LIGLSTLKIR *TYAV*TPFR RHLNAQ*

```

m088/a088 99.5% identity over a 205 aa overlap

	150	160	170	180	190	200
m088.pep	XXXXXXXXXXXXXXXXXXXXXXXXXGVVGFVLVSYLTIVGTSNAVNLTDLGLDGLA					
a088	IAGLALFYLAANSANNILIVPFFKQIALPLGVVGFVLVSYLTIVGTSNAVNLTDLGLDGLA					



283

	150	160	170	180	190	200
	210	220	230	240	250	260
m088.pep	TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA					
a088	TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA					
	210	220	230	240	250	260
	270	280	290	300	310	320
m088.pep	QVFMGVDVGALALGAALGTVAIVIRQEFVLVIMGGLFVVEAVSVMLQVGWYKTKKRIFLM					
a088	QVFMGVDVGALALGAALGTVAIVIRQEFVLVIMGGLFVVEAVSVMLQVGWYKTKKRIFLM					
	270	280	290	300	310	320
	330	340	350	360	370	
m088.pep	APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRLHNAQX					
a088	APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRLHNAQX					
	330	340	350	360	370	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 289>:

```

g089.seq
1  ATGCCGCCCA AAATCACGAA GAGCGGGTTT TGCAAACCGG CAATCGCGGC
51  GGCGGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATG AATACCACGC
101 CGTTTTTCTC GCCGATTTTT TCCACACGGT GCGGCAAGCC TTGGAAGGTT
151 TTGACGTGTT CCAGCAATGC TTCGCGCGGC AAACCGACGG CCTCGCACAA
201 AGCCACGGCA GCCATAACGT TGGCGGCGTT GTGCAAACCT TGCAGCGGGA
251 TGTCTTGCGT AGAAATCAAA TCTTCATTGC CTTGTTTTAA ACAGCCCGTC
301 CCGCGTTCCA ACCAAAAATC GGCTTCGTGT TCCAAGGAAA ACCGTTTCAC
351 TTCACGCCCT GCCCGTTTCA TGGCGCGGCA GAACACGTCG TCCGCATTCA
401 AAACCTGCAC TCCATCGCCA CGGAAAATCT CGGCTTTGGT ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 290; ORF 089.ng>:

```

g089.pep
1  MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFSPIF STRCGKPWKV
51  LTCSSNASRG KPTASHKATA AITLAALCKP CSGMSCVEIK SSLPCFKQPV
101 PRSNQKSASC SKENRFTSRP ARFMARQNTS SAFKTCTPSP RKISALVCA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 291>:

```

m089.seq
1  ATGCCGCCCA AAATCACKAw GAGCGGATTT TGCAAACCGG CAATCGCGGC
51  GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA AACACCACGC
101 CGTTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC TTGGAAGGTT
151 TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG CCTCACACAA
201 AGCCACKGCA GCCATGACGT TAGCGGCGTT GTGCAKACCT TGCAACGGwA
251 TGTCTTGCGT GACAATCAAA TCTTCATTGC CTTGTTTCAG GCGGCCTGTC
301 TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA ACCATTTTAC
351 CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG TCCGCATTCA
401 AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCATAG

```

This corresponds to the amino acid sequence <SEQ ID 292; ORF 089>:

```

m089.pep
1  MPPKITXSGF CKPAIAAAVA PTFVPLLSSI NTTPFFSPIF STRCGRPWKV
51  LTCSSNASRD KPMASHKATA AMTLAALCXP CNGMSCVTIK SSLPCFRRPV
101 SRSNQKSASC SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 089 shows 88.6% identity over a 149 aa overlap with a predicted ORF (ORF 089.ng) from *N. gonorrhoeae*:

m089/g089

	10	20	30	40	50	60
m089.pep	MPPKITXSGFCKPAIAAAVAPTFVPLLSSINTTPFFSPIFSTRCGRPWKVLTCSSNASRD					
g089	MPPKITXSGFCKPAIAAAVAPTFVPLLSSMNTTPFFSPIFSTRCGRPWKVLTCSSNASRG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m089.pep	KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRSNQKSASCSNENHFTSRP					
g089	KPTASHKATAAITLAALCKPCSGMSCVEIKSSLPCFKQPVPRSNQKSASCSKENRFTSRP					
	70	80	90	100	110	120
	130	140	150			
m089.pep	ARFIARQNASSAFKTCTPSPRKILALVCAX					
g089	ARFMARQNTSSAFKTCTPSPRKISALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 293>:

## a089.seq

```

1  ATGCCGCCTA AAATCACGAA GAGCGGATT TGCAAACCGG CAATCGCGGC
51  GCGGTCGCA CCGACGTTCG TGCCTTTGCT GTCGTCGATG AACACCACGC
101 CATTTTTCTC GCCGATTTT TCCACGCGGT GCGGCAGGCC TTGAAAGGTT
151 TTGACGTGTT CGAGCAATGC TTCGCGCGGC AAACCGACGG CTTCGCACAA
201 GGCAACGGCA GCCATCACGT TAGTGCGGTT GTGCAAGCCT TGCAGCGGAA
251 TATCTTGCGT GGCAATCAAA TCTTCATTGC CTTGTTTCAG GCGACCTGTC
301 TCACGTCCA ACCAAAATC GGCTTCGTAT TCCAACGAAA ACCATTTCAC
351 CTCGCGCCCG GCGCGCTTCA TCGCACGACA GAACGCATCG TCCGCATTCA
401 AAACCTGCAC ACCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 294; ORF 089.a>:

## a089.pep

```

1  MPPKITXSGF CKPAIAAAVA PTFVPLLSSM NTPFFSPIF STRCGRP*KV
51  LTCSSNASRG KPTASHKATA AITLVALCKP CSGISCVAIK SSLPCFRRPV
101 SRSNQKSASY SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

```

m089/a089 91.9% identity over a 149 aa overlap

	10	20	30	40	50	60
m089.pep	MPPKITXSGFCKPAIAAAVAPTFVPLLSSINTTPFFSPIFSTRCGRPWKVLTCSSNASRD					
a089	MPPKITXSGFCKPAIAAAVAPTFVPLLSSMNTTPFFSPIFSTRCGRPXKVLTCSSNASRG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m089.pep	KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRSNQKSASCSNENHFTSRP					
a089	KPTASHKATAAITLVALCKPCSGISCVAIKSSLPCFRRPVSRSNQKSASYSNENHFTSRP					
	70	80	90	100	110	120
	130	140	150			
m089.pep	ARFIARQNASSAFKTCTPSPRKILALVCAX					
a089	ARFIARQNASSAFKTCTPSPRKILALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 295>:

## g090.seq

```

1  ATGCGCGTAG TCGAGCAAAT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA

```

```

51  TGTTTCATCAC CGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
101 TGGAAGCTGG AAAGCTCcca CACCCACACG TCCGCCTTTT TGCCTTCgcg
151 ctgCAATtct gcctccaaga cgggcgtacc gatATTGCCC GCAATGAcgg
201 tatccagccc gcacttgatg CAGAGatagc ggaccaggct ggttaccgTG
251 GTTttgccgt tgctgCcggt aatcgCaatc accttgtcgC CGCGGCGGtt
301 cAcaaTGTCc gccagCAATt ggATGTCGCC TAgCACGCGC .ccgcgTtT
351 TGCTtga

```

This corresponds to the amino acid sequence <SEQ ID 296; ORF 090.ng>:

```

g090.pep
1  MRVVEQIVVA VEMVFGNVHH RRRSRAQAFG VFQLEAGKLP HPHVRLFAFA
51  LQFCLQDGR TDIARN DGIQ P ALDAE IADQA GYRGFAVAAG NRNHLVAAA V
101 HNV RQQLDVA XHAXRRFA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 297>:

```

m090.seq
1  ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
51  TGTTTCAGCAC CGTCGCCGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAGT
101 TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
151 CTGCCATTCC GCCTCCAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
201 TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
251 GTTTTGCCGT TGCTGCCGGT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
301 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT .CCGCCGTTT
351 TGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 298; ORF 090>:

```

m090.pep
1  MRIVEQVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
51  LPFRLQNRRA DIARN DGIQ P ALDTE IADQA RYRGFAVAAG NRNYLVVPAV
101 HNV RQQFDVA QHAXRRFA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 090 shows 83.9% identity over a 118 aa overlap with a predicted ORF (ORF 090.ng) from *N. gonorrhoeae*:

m090/g090

	10	20	30	40	50	60
m090.pep	MRIVEQVVAVEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRA					
g090	MRVVEQIVVAVEMVFGNVHHRRRSRAQAFGVFQLEAGKLPHPHVRLFAFALQFCLQDGR T					
	10	20	30	40	50	60
	70	80	90	100	110	119
m090.pep	DIARN DGIQ PALDTEIADQARYRGFAVAAGNRNYLVVPAVHNV RQQFDVAQHAXRRFAX					
g090	DIARN DGIQ PALDAE IADQAGYRGFAVAAGNRNHLVAAA VHNVRQQLDVAXHAXRRFAX					
	70	80	90	100	110	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 299>:

```

a090.seq
1  ATGCGCGTAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
51  TGTTTCAGCAC TGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
101 TGGAAACTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
151 CTGCAATTCC GCCTCCAAA CCGGCGCGCC GATATTGCCC GCGATAACGG
201 TATCCAGCCC AACTTGTATG CAGAGATAGC CGACCAGGCT CGTTACCGTG
251 GTTTTGCCGT TGCTGCCGGT AATCGCAATC ACCTTGTCGC CGCGGCGGTT
301 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT C.CGCCGTTT
351 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 300; ORF 090.a>:

## a090.pep

1 MRVVEQVVVA VEMVFGNVQH CRRSRAQAFG VFQLETGKLQ HPHVRLFAFA  
 51 LQFRLQNRRA DIARDNGIQP TLDAEIADQA RYRGFAVAAG NRNHLVAAAV  
 101 HNVRRQFDVA QHAXRRFA\*

m09/a090 91.5% identity over a 117 aa overlap

	10	20	30	40	50	60
m090.pep	MRIVEQVVVA	VEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFR	LQNRRA			
a090	MRVVEQVVVA	VEMVFGNVQHRRRSRAQAFGVFQLETGKLQHPHVRLFAFALQFRLQNRRA				
	10	20	30	40	50	60

	70	80	90	100	110	119
m090.pep	DIARDNGIQPALDTEIADQARYRGFAVAAGNRNRYLVVPAVHNVRRQFDVAQHAXRRFAX					
a090	DIARDNGIQPTLDAEIADQARYRGFAVAAGNRNHLVAAAVHNVRRQFDVAQHAXRRFAX					
	70	80	90	100	110	

The following partial DNA sequence was identified in *N. gonorrhoeae*

g090-1.seq This sequence contains multiple stop codons (not shown)

This corresponds to the amino acid sequence < ORF 090-1.ng>:

g090-1.pep (not shown)

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 2>:

## m090-1.seq

1 ATGACGGCGT TTGCATTTC GACGGCATCA CAAAGCCTTA AACGCTTCGA  
 51 TAAACACTTC CGAACGGTGC GCGTAGCCTT TGAACATATC AAAGCTCGCG  
 101 CAGGCGGGG TGAGCAACAC AATATCGCCT GCTTCGGCTT GGGCATATGC  
 151 CGTCTGAACG GCTTCTCCCA AAGTGGCGCA GTCGGTCATA TTCAAGCCGC  
 201 AGCCGTCCAA ATCGCGGCGG ATTTGCGGCG CATCGACACC AATCAAGAAC  
 251 ACGCCTTTTG CTTGCTTAC CAGTGCATCG CGCAGGGGCG TGAAGTCCTG  
 301 CCCTTTACCC ATGCCGCCCA AAATCACGAA GAGCGGATTT TGCAAACCGG  
 351 CAATCGCGGC GGCAGTCGCG CCGACATTCTG TGCCTTGTCT GTCGTCGATA  
 401 AACACCACGC CGTTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC  
 451 TTGGAAGGTT TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG  
 501 CCTCACACAA AGCCACGGCA GCCATGACGT TAGCGCGCTT GTGCAGACCT  
 551 TGCAACGGAA TGTCTTGCCT GACAATCAAA TCTTCATTGC CTTGTTTCAG  
 601 GCGGCTGTCT TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA  
 651 ACCATTTTAC CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG  
 701 TCCGCATTCA AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT  
 751 ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA  
 801 TGTTGAGCAC CGTCGCCGCA GTCGGACGCA GGCCTTTCGT GTTTTCCAGT  
 851 TGGAAAGCTGG AAAGCTCCAA CACCCACAG TCCGCCTTTT TGCTTCCGCG  
 901 CTGCCATTCC GCCTCCAAA CCGCGGTGCC GATATTGCC GCGATAACGG  
 951 TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG  
 1001 GTTTTGCCGT TGTGCGCGT AATCGCAATT ACCTTGTCTG CCCGGCGGTT  
 1051 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT CCGCCGTTTT  
 1101 GCTTGAACGC CTCAATATCC GGCTGCCGCT CGCTGATGCC GGGACTGAGA  
 1151 GCCAGAATAT CGAAACCGTT GTCCAGCGCA TCTTTCAGAC GGCCCGTGTA  
 1201 AAACACCAAC CCGTCAAACA TCTTACCGAT TTGCGACACG CGTTCCGGCT  
 1251 TCAGCTCCGC ATCATACGCA GCAACCTCCG CGCCGTTTTT GCGCAGGTAG  
 1301 GCAATCATGG AAATACCGT ACCGCGAGT CCGGCGACGA GGATTTTTTT  
 1351 GTTTTGAAAA GTCATTTTGG TTTGTCCTAA

This corresponds to the amino acid sequence <SEQ ID 3; ORF 090-1>:

## m090-1.pep

1 MTAFAFQTAS QSLKRFDKHF RTVRVAFEDI KARAGGAEQH NIACFGLGIC  
 51 RINGFSQSGA VGHIAAAVQ IAADLRIDT NQEHAFCLAY QCIAQGREVL  
 101 PETHAAQNHE ERILOTGNRG GSRADIRAFV VVDKHHAVFL ADFFHAVRQA  
 151 LEGFDVFEQC FARQTDGLTQ SHGSHDVSGV VQTLQRNVLR DNQIFIALFO  
 201 AACLAQFQEI SFVFQRKPFY LAPGTLHRAA ERIVRIQNLH AVATENLGFQ  
 251 MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA  
 301 LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRNRYLVVPAV  
 351 HNVRRQFDVA QHASAVLLER LNIRLPLADA GTESQNIETV VQRIQFQARV  
 401 KHQPVKHLTD LRHAFRLQLR IIRSNLRAVF AQVGNHGNTR TAESGDEDDF  
 451 VLKSHFGLS\*

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 303>:

g091.seq

```

1  ATGGAAATAC CCGTGCCGCC AAGTCCGGCG ACGAGGATTT TTTTGTGTTGA
51  AAGTCATTTT GGTTTTGTCC TAAAACAAAT CATATTGGGC AGGAGACGTC
101 CGCCCTTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCCG ATTAATAACC
151 CGCCTTCAGG CGTTGGTCAT TGTCGCAGCT GTTTTGGTCT CCGTTTGTGAC
201 AAGCCTTGCC AAGCCATTGT TGAGCGAGCG CAAGGTCTTG GCGCACGCCG
251 CGTCCATCGT AATACATCAA GCCCAAATTG TATTGGGCTT GGGCATCCCC
301 TTGTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 304; ORF 091.ng>:

g091.pep

```

1  MEIPVPPSPA TRIFLFESHF GFVLKQIILG RRRPPLPKPL SDGIASRLIT
51  RLQALVIVAA VLVSVLTSIA KPLLSEKVL  AHAASIVIHQ AQIVLGLGIP
101  LF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 305>:

m091.seq

```

1  ATGGAAATAC CCGTACCGCC GAGTCCGGCG ACGAGGATTT TTTTGTGTTGA
51  AAAGTCATTT TGGTTTGTCC TAAAACAAAT CATATTGAGC AGGAGATGTC
101 CGCCCTTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
151 CGCCTTCAGG CGTTGGTCAT TGTCGCAGCC GTCTTGGTCT CCGTTTGTGAC
201 AAGCCTTGCC AAACCATCTT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG
251 CGTCTTTCGG CATAATCAC  GCCCAAATTG TTTTGGGCTT GGGCTACCCC
301 CTGCGC...

```

This corresponds to the amino acid sequence <SEQ ID 306; ORF 091>:

m091.pep

```

1  MEIPVPPSPA TRIFLFEKSF WFLVKQIILS RRCPPPLPKPL SDGIASCSIT
51  RLQALVIVAA VLVSVLTSIA KPFLCKGAVL  AHAASFGIHH AQIVLGLGYP
101  LR.

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 091 shows 84.2% identity over a 101 aa overlap with a predicted ORF (ORF 091.ng) from *N. gonorrhoeae*:

m091/g091

	10	20	30	40	50	60
m091.pep	MEIPVPPSPATRIFLFEKSFVFLVKQIILSRRCPPLPKPLSDGIASCSITRLQALVIVAA					
	:       :					
g091	MEIPVPPSPATRIFLFESHFGFVLKQIILGRRRPPLPKPLSDGIASRLITRLQALVIVAA					
	10	20	30	40	50	60
	70	80	90	100		
m091.pep	VLVSVLTSIAKPFLCKGAVLAHAASFGIHHQAIVLGLGYPLR					
	:  :      :   :					
g091	VLVSVLTSIAKPLLSEKVLHAASIVIHQAIVLGLGIPLFX					
	70	80	90	100		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 307>:

a091.seq

```

1  ATGGAAATAC CCGTGCCGCC AAGTCCGGCG ACGAGGATTT TTTTGTGTTG
51  GAAATCATTT TGGTTTGTCC TAAAACAAAT CATATTGAGC AGGGGATGTC
101 TGATCCTGCT CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
151 CGCTTTCAGG CGTTGGTCAT TGTCGCAGCT GTCTTGGTAT CCGTTTGTGAC
201 AAGCCTTGCC AAGCCATTCT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG
251 CGTCTTTCGG CATAATCAC  GCCCAAATTG TTTTGGGC

```

This corresponds to the amino acid sequence <SEQ ID 308; ORF 091.a>:

a091.pep

```

1  MEIPVPPSPA TRIFLFWKSF WFLVKQIILS RGCLILLKPL SDGIASCSIT
51 RFQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLG

```

m091/a091 93.8% identity over a 96 aa overlap

```

              10      20      30      40      50      60
m091.pep    MEIPVPPSPATRIFLFEKSFWEVLVKQIILSRRCPPLEKPLSDGIASCSITRLQALVIVAA
              |||||
a091         MEIPVPPSPATRIFLFWKSFWEVLVKQIILSRGCLILLKPLSDGIASCSITRFQALVIVAA
              10      20      30      40      50      60

              70      80      90     100
m091.pep    VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAAQIVLGLGYPLR
              |||||
a091         VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAAQIVLG
              70      80      90

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 309>:

g092.seq

```

1  ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGTGCGC
51  AAACGGTCAG ACCTTTAAAA TAACGCCTTT ACGCACTAAA AACCACCCGG
101 AACGCAACAT TATGATGAAA AATCGAGTAA GCAACATCCA TTTTGTCCGGT
151 ATCGGCGGCG TCGGCATGAG CCGTATCGCC GAAGTCTTGC ACAATTTGGG
201 CTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
251 TGAGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT
301 AACGGTgcgg ATGTCGTCGT TGCCTCTACC GCCGTCAAGA AAGAAaatcC
351 CGAAGTtgtc gcTGC GTTGG AGCGGCAAAT TCCCGTTATT CCGCGCGCCT
401 TGATGCTGGC AGAGCTGATG CGCTTCCGTG ACGgcatcgc cattgccgT
451 ACGCACGGCA AAACACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTC GACCCCACTT TCGTTATCGG CGGCAAACTC AACGCCGCGC
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAATCCGATG CCTCTTTCCT ACATCTGACC CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTTGT GTGTTGACAG CGAACACGTC CGCGCGATT TGCCTCAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GGCGCGCAA TGAATTAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGc gtggcGCTgg
1001 aagtcGgCGC ATcggttgAA GCGAtcCAA AaggCTTGCT CGGCTTTGAA
1051 GCGCTCGGCC GCGCTTCCA AAAATAcggc gacatCAagt tgccaaacgg
1101 cggGaccgCT TTgctGGTGG ACGATTAcgg ACACCACCC GTCGAAATGG
1151 CGGcaaccct tgccgcTGCA CGCGGCGCGT ATCCGAAAA acgtTTGGTG
1201 CtgcCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTACTCAATA CCGTTGatgc GCTGGTACTG ACCGAAGTTT
1301 AtgccgcccG CGAAGAGCCG GTTGCCGCCG CCGactcCCG CGCCTTGGCG
1351 CGTGCTATCC GCGTATTGGG CAAACTTGAG CCGATTTACT GCGAAAatgt
1401 cgccgACCTG CCGCAAATGC TGATGAATGT TTTACAGGAT Ggcgatgttg
1451 tggtgAATAT GggTgcggga agcatcaacc gcgttccttc cgcgctgttg
1501 gaattgtcga AACAGAttg A

```

This corresponds to the amino acid sequence <SEQ ID 310; ORF 092.ng>:

g092.pep

```

1  MFFISIRYIF VRKLWCANGQ TFKITPLRTK NQPERNIMMK NRVSNIHFVG
51  IGGVMSGIA EVLHNLGFKV SGSDQARNAA TEHLSSLGIQ VYPGHTAEHV
101 NGADVAVAST AVKKENPEVV AALERQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK

```

```

251 AFLCVDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351 GVRRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP VAAADSRALA
451 RAIRVLGKLE PIYCENVADL PQMLMNVLQD GDVVLNMGAG SINRVPSALL
501 ELSKQI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 311>:

m092.seq

```

1 ATGTTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
51 AAACGGTCAG CCCTTTAAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTTA CCAACATCCA TTTTGTGCGT
151 ATCGGCGGCG TCGGCATGAG CGGCATCGCC GAAGTCTTGC ACAATTTGGG
201 CTTTAAAGTT TCCGGTTCGG ATCagGCGCG AAATGCCGCT ACCGAGCATT
251 TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC CGAACACGTT
301 AACGGTGCGG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGCG
451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCGC
551 GCACTAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTTGT GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCAAA TGAAATCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GCGTCGGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 CGGGACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CCGCGACCCCT TGCCGCCGCA CGCGGCGCGT ATCTGGAAAA ACGTTTGGTA
1201 CTCGCTTCC AGCCGACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAA GTCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CCGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTGGG CAAACTCGAG CCGATTACT GCGAAAACGT
1401 TGCCGATCTG CCGGAAATGC TGTGAACGT TTTGCAGGAC GGCGACATCG
1451 TGTTGAATAT GGGCGCGGGA AGCATCAACC GCGTCCCCGC CGCGCTGCTG
1501 GCATTGTCTGA AACAGATTTG A

```

This corresponds to the amino acid sequence <SEQ ID 312; ORF 092>:

m092.pep

```

1 MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIHFVG
51 IGGVMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHAEHV
101 NGADVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351 GVRRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYLEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451 RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMGAG SINRVPAALL
501 ALSKQI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 092 shows 96.6% identity over a 506 aa overlap with a predicted ORF (ORF 092.ng) from *N. gonorrhoeae*:

m092/g092

```

10          20          30          40          50          60
m092.pep   MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVMSGIA

```

```

g092      ||||| ||||| ||||| ||||| : ||||| ||||| ||||| ||||| |||||
          10      20      30      40      50      60
          70      80      90     100     110     120
m092.pep  EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVGADVVVTSTAVKKENPEVV
          ||||| ||||| ||||| ||||| : ||||| ||||| ||||| ||||| |||||
g092      EVLHNLGFKVSGSDQARNAATEHLSSLGIQVYPGHTAEHVGADVVVASTAVKKENPEVV
          70      80      90     100     110     120
          130     140     150     160     170     180
m092.pep  AALEQQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
          ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g092      AALERQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL
          130     140     150     160     170     180
          190     200     210     220     230     240
m092.pep  NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g092      NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID
          190     200     210     220     230     240
          250     260     270     280     290     300
m092.pep  FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKTFTVHV
          ||||| ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g092      FIHRMPFYGKAFLCVDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKTFTVHV
          250     260     270     280     290     300
          310     320     330     340     350     360
m092.pep  QMGHEQGSFEVVLNMPGRHNVNALAAIGVALEV GASVEAIQKLLGFEGVGRRFQKYG
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g092      QMGHEQGSFEVVLNMPGRHNVNALAAIGVALEV GASVEAIQKLLGFEGVGRRFQKYG
          310     320     330     340     350     360
          370     380     390     400     410     420
m092.pep  DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLA FQPHRYTRTRDLFEDFTK
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g092      DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYPEKRLVLA FQPHRYTRTRDLFEDFTK
          370     380     390     400     410     420
          430     440     450     460     470     480
m092.pep  VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIY CENVADLP EMLLNVLQD
          ||||| ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g092      VLNTVDALVLTEVYAAGEEPVAAADSRALARAIRVLGKLEPIY CENVADLP QMLMNVLQD
          430     440     450     460     470     480
          490     500
m092.pep  GDIVLNMGAGSINRVPAALLALSKQIX
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g092      GDVVLNMGAGSINRVPSALLELSKQIX
          490     500

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 313>:

a092.seq

```

1  ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
51 AAACGGTCAG CCCTTTTAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTGA CCAACATCCA TTTGTCTGGT
151 ATCGGCGCGC TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTTGGG
201 TTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
251 TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT

```



```
301 AACGGTGCGG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGT GTTGAGAAGC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGTG GTATTGACAG CGAACACGTC CGCGCGATT TGTCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGGG CGATACCGCC GACATCTACG
851 CCACCACAT CGAAAACGTC GGCGCGCAAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GGTGTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 TGGAAACGCG CTCTTGTTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CGGCGACCTT TTCCGCCGCA CGCGGCGCGT ATCCGGAAAA ACGTTTGGTA
1201 CTCGCCCTTC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CTGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTGGG CAAACTCGAG CCGATTTACT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC TGTGTAACGT TTTGCAGGAC GGCGACATCG
1451 TGTGAATAT GGGTGCGGGA AGCATCAACC GCGTCCCCGC CGCGCTGCTG
1501 GAATTGTCTG AACAGATTG A
```

This corresponds to the amino acid sequence <SEQ ID 314; ORF 092.a>:

**a092.pep**

```
1 MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIHFGV
51 IGGVGMMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
101 NGADVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
351 GVGRRFQKYG DIKL PNGGTA LLVDDYGHHP VEMAATLSAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451 RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMGAG SINRVPAALL
501 ELSKQI*
```

**m092/a092** 99.4% identity over a 506 aa overlap

	10	20	30	40	50	60
m092.pep	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMMSGIA					
a092	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMMSGIA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m092.pep	EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKENPEVV					
a092	EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKENPEVV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m092.pep	AALEQQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL					
a092	AALEQQIPVIPRALMLAELMRFRDGIAGTHGKTTTTSLTASILGAAGLDPTFVIGGKL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m092.pep	NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					
a092	NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					

292

	190	200	210	220	230	240
	250	260	270	280	290	300
m092.pep	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV					
a092	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFTVHV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m092.pep	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKLLGFEGVGRRFQKYG					
a092	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKLLGFEGVGRRFQKYG					
	310	320	330	340	350	360
	370	380	390	400	410	420
m092.pep	DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLA FQPHRYTRTRDLFEDFTK					
a092	DIKLPNGGTALLVDDYGHHPVEMAATLSAARGAYPEKRLVLA FQPHRYTRTRDLFEDFTK					
	370	380	390	400	410	420
	430	440	450	460	470	480
m092.pep	VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
a092	VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
	430	440	450	460	470	480
	490	500				
m092.pep	GDIVLNMGAGSINRVPAALLALSQIX					
a092	GDIVLNMGAGSINRVPAALLELSQIX					
	490	500				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 315>:

g093.seq

```

1  aTGCAGAATt ttgGCAAAGT ggccgtATTG ATGGGtggtT TTTCCAGCGA
51  ACGAGAAatc tcgcTGGACA GcGgTACCGC CATTTTGAAC GCCTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGACC CTAAGGAAAC GCCGTTATCC
151 GAACTGAAGG AGCGGGGCTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAAC TGGGGCATT
251 CCTATACCGG CAGCGGTGTC GCCGCTCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTGTGCA GGCATTGGGA TTACCCGTC CCGAGTTCGC
351 CGTACTGTAC GATGATACCG ATTTTCGATG CGTCGAAGAA AAATTGGGTC
401 TGCCGATGTT TGTGAAGCCG GCGGCCGAAG GCAGCAGCgt cggcgtggt
451 aAAGTCAAAG AAAaaggccg TCTGAAAAGC GTTtacgaag aatTGAAaCA
501 CCTTcagggg cgaAatcatt gccgAacgTT TTATCGGCGG CGGCGAATAT
551 TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATA ACATCATCCC
601 CGCAACCGAG TTTTACGAct acgaagccaa GtacaacCGA GACGAcacca
651 tttatCAATG TCCTTCGGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTCGCGG CGCACAGGCA ATCGGTGCGG AAGGTCGCGT
751 GCGCGTCGAT TTCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
801 TCAACACCTT GCCCGGTATG ACCGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 316; ORF 093.ng>:

g093.pep

```

1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKERGFQTA FNILHGTYGE DGAVQGAELE LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLV DDTDFDAVEE KLGLPMFVKP AAEGLSSGVV
151 KVKEKGR LKS VYEELKHLQG RNHCRTFYRR RRIFLPRPER QRAARHTHP
201 RNRVLR LRSQ VQPRRHLSM SFGRFDRSRR KPDARTGGS RTGNRCGLRL
251 ARRFPQRYR QTLVGNQHP ARYDRP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 317>:

m093.seq

```

1  ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGATC CTAAAGAAAC CCCATTGTCT
151 GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCrAA GACGGGGCGG TTCAGGGTGC ATTGGAACTG TTGGGCATTC
251 CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351 CGTCCTGCAC GACGACACTG ATTTTCGATC CGTCGAAGAA AAATTGGGCC
401 TGCCGATGTT TGTGAAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451 AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTTACGAAG AATTGAAACA
501 CCTTCAGGG. CGAAATCATT GCCGAACGTT TTATCGGCGG CGGCGAATAT
551 TCCTGCCCCG TCCTGAACCG CAAAGGGCTG CCCGGCATAC ACATCATTC
601 CGCAACCGAG TTTTACGACT ACGAAGCCAA GTACAACCGC GACGACACCA
651 TTTATCAATG TCCTTCGGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTCGCGG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
801 TCAACACCCT GCCCGGTATG ACGAGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 318; ORF 093>:

```

m093 . pep
1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKAQGFQTA FNILHGTYGX DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
151 KVKGKGR LKS VYEELKHLQX RNHCRTFYRR RRIFLPRPER QRAARHTHS
201 RNRVLR LRSQ VQPRRHLSM SFGRFDRSR KPDARTGGS RAGNRCGR LR
251 ARRFPQRYRR QTLSVGNQHP ARYDEP*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 093 shows 96.7% identity over a 276 aa overlap with a predicted ORF (ORF 093.ng) from *N. gonorrhoeae*:

m093/g093

	10	20	30	40	50	60
m093 . pep	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLSELKAQGFQTA					
g093	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLSELKERGFQTA					
	10	20	30	40	50	60
m093 . pep	FNILHGTYGX DGAVQGALELLGI PYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH					
g093	FNILHGTYGEDGAVQGALELLGI PYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLY					
	70	80	90	100	110	120
m093 . pep	DDTDFDAVEEKLGLPMFVKPAAEGSSVGVV KVKGKGR LKSVYEELKHLQXRNHCRTFYRR					
g093	DDTDFDAVEEKLGLPMFVKPAAEGSSVGVV KVEKGR LKSVYEELKHLQGRNHCRTFYRR					
	130	140	150	160	170	180
m093 . pep	RRIFLPRPERQRAARH <sub>T</sub> HSRNRVLR LRSQVQPRRHLSMSFGRFDRSR RKP <sub>D</sub> ARTGGS <sub>R</sub>					
g093	RRIFLPRPERQRAARH <sub>T</sub> HPNRVLR LRSQVQPRRHLSMSFGRFDRSR RKP <sub>D</sub> ARTGGS <sub>R</sub>					
	190	200	210	220	230	240
m093 . pep	RAGNRCGR L <sub>R</sub> ARRF <sub>P</sub> QRYRRQTL <sub>S</sub> VGNQHPARYDEP <sub>X</sub>					
g093	RTGNRCGR L <sub>R</sub> ARRF <sub>P</sub> QRYRRQTL <sub>S</sub> VGNQHPARYDR <sub>PX</sub>					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 319>:

```
a093.seq
1  ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGATC CCAAGGAAAC CCCATTGTCT
151 GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAAGT TTGGGCATTG
251 CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351 CGTCCTGCAC GACGACACTG ATTTTCGATG CGTCGAAGAA AAATTGGGCC
401 TGCCGATGTT TGTGAAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451 AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTACGAAG AATTGAAACA
501 CTTTCAGGG. CGAAATCATT GCCGAACGGT TTATCGGCGG CCGCGAATAT
551 TCCTGCCCTG TGTGAACGG CAAAGGCCTG CCCGGCATA ACATCATCCC
601 CGCGACCGAG TTTTATGACT ACGAAGCCAA GTACAACCGC AACGACACCA
651 TTTATCAATG TCCTTCGGAA GATCTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTGCGCG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGAAA
801 TCAACACCCT GCCCGGTATG ACCGGCCATA G
```

This corresponds to the amino acid sequence <SEQ ID 320; ORF 093.a>:

```
a093.pep
1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKAQGFQTA FNILHGTYGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGSSVGUV
151 KVKGKGR LKS VYEELKHFX RNHCRTVYRR RRIFLPCVER QRPARTHHP
201 RDRV L* LRSQ VQQRHLSM SFGSRDSRR KPDARTGGS RAGNRCGR LR
251 ARRF PQRYRR QTLVGNQHP ARYDRP*
```

m093/a093 95.7% identity over a 276 aa overlap

```

      10      20      30      40      50      60
m093.pep  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLSELKAQGFQTA
a093      |||
      10      20      30      40      50      60
m093.pep  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLSELKAQGFQTA
a093      |||

      70      80      90     100     110     120
m093.pep  FNILHGTYGDGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH
a093      |||
      70      80      90     100     110     120
m093.pep  FNILHGTYGDGAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH
a093      |||

     130     140     150     160     170     180
m093.pep  DDTDFDAVEEKLGLPMFVKPAAEGSSVGUVKVKGKGR LKS VYEELKHLQXRNHCRTFYRR
a093      |||
     130     140     150     160     170     180
m093.pep  DDTDFDAVEEKLGLPMFVKPAAEGSSVGUVKVKGKGR LKS VYEELKHFXRNHCRTVYRR
a093      |||

     190     200     210     220     230     240
m093.pep  RRIFLPRPERQRAARHTHHSRNRVLR LRSQVQPRRHLSMSFGRFDRSRRKPDARTGGS R
a093      |||
     190     200     210     220     230     240
m093.pep  RRIFLPCVERQRPARTHHPRDRV LXLRSQVQQRHLSMSFGRSDRSRRKPDARTGGS R
a093      |||

     250     260     270
m093.pep  RAGNRCGR LRARRFPQRYRRQTLVGNQHPARYDEPX
a093      |||
     250     260     270
m093.pep  RAGNRCGR LRARRFPQRYRRQTLVGNQHPARYDRPX
a093      |||
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 321>:

```
g094.seq
1   ATGTATTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
51  GCCGCCGATA ACGAAAGTGG GGTTCGAGTCC TGCCGCGCCG AGGATGGAGG
101 CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTAccggc aatggcgatg
151 cCGTCACGGA AGCGCATCAG CTCTGCCAGC ATCAAGGCGC GCGGAATAAC
201 GGGAAATTTGC CGTCCAACG CAgcgacaAC TTCGGgattT TCTTTCTTGA
251 CGGCGGTAGA GGCAACGACG ACATccgcAC CGTTAACGTG TTCTGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 322; ORF 094.ng>:

```
g094.pep
1   MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
51  PSRKRISSAS IKARGITGIC RSNAATTSGF SFLTAVEATT TSAPLTCSAV
101 WPG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 323>:

```
m094.seq
1   ATGTATTCGC CTTTGCCCAA GCGGGCGTTA GTGCCTGCGG CGTTGAGTTT
51  GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
101 CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
201 GGGAAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATTT TCTTTTTTGA
251 CGGCGGTAGA GGTAAACGAC ACATCCGCAC CGTTAACGTG TTCGGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 324; ORF 094>:

```
m094.pep
1   MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
51  PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
101 WPG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 094 shows 95.1% identity over a 103 aa overlap with a predicted ORF (ORF 094.ng) from *N. gonorrhoeae*:

```
m094/g094

      10      20      30      40      50      60
m094.pep  MYSPLPKRALVPAALSLPPIITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRIINSAN
          |||
g094      MYSPLPKRALVPAALSLPPIITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRISSAS
          |||

      70      80      90     100
m094.pep  IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
          |
g094      IKARGITGICRSNAATTSGFSFLTAVEATTTSAPLTCSAVWPGX
          |

      70      80      90     100
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 325>:

```
a094.seq
1   ATGTATTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
51  GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
101 CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
201 GGGAAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATTT TCTTTTTTGA
251 CGGCGGTAGA GGTAAACGAC ACATCCGCAC CGTTAACGTG TTCTGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 326; ORF 094.a>:

**a094.pep**

```

1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
51  PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
101 WPG*

```

**m094/a094** 100.0% identity over a 103 aa overlap

```

              10      20      30      40      50      60
m094.pep     MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
              |||||
a094          MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
              10      20      30      40      50      60

              70      80      90     100
m094.pep     IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
              |||||
a094          IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
              70      80      90     100

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 327>:

```

g095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151 AACACACAAA AAGGCTTTGC CGTAGAAGGG CACACGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGGGTCA GTGTAGGAAA
301 GAGGCATCGG ATCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351 CGGGCGTTGG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 328; ORF 095.ng>:

```

g095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRGQCRK
101 EASDRRLRQR CIRLCPSGRW CLRR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 329>:

```

m095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CACACGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
301 GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351 CGGGCGTTAG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 330; ORF 095>:

```

m095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRGQCRK
101 DASDRRLRQR CIRLCPSGRX CLRR*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 095 shows 97.6% identity over a 124 aa overlap with a predicted ORF (ORF 095.ng) from *N. gonorrhoeae*:

```

m095/g095
              10      20      30      40      50      60
m095.pep     MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
              |||||

```

```

g095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           10      20      30      40      50      60

           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
g095      HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRGQCRKEASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
g095      CLRRX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 331>:

```

a095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTCG CCGCCGACGT
51 TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAATC GCGCGGACGT GTTCGCTGTC
151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CATA CGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGCTTC TCAACACTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
301 GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351 CGGGCGTTGG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 332; ORF 095.a>:

```

a095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51 NTQKGFAVEG HTVDEIDKRL MQLLNTVPVG IHMFVDIGN DGHNRCQCRK
101 DASDRRLRQR CIRLCPSGRW CLRR*

```

m095/a095 96.0% identity in 124 aa overlap

```

           10      20      30      40      50      60
m095.pep  MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           |||||
a095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           10      20      30      40      50      60

           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
a095      HTVDEIDKRLMQLLNTVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
a095      CLRRX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 333>:

```

g096.seq
1  ATGGCCGGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51 CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
101 GCCTGTGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTGGGCTTC
151 GGTCAAATCT TCCGAAGGAC ATTGAataat ggtgTCGTCT CGGttgtaCt

```

```

201  tggcttcgta gTCGTAAAC TCGGTTGCGG GGATGATGTG TATGCCGGGC
251  AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301  AcgtTcggca atgaTTtcgc ccctgAAGGT GttTCAattc ttcgtaAACG
351  CTTTTCAGAC ggccttTTTC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 334; ORF 096.ng>:

## g096.pep

```

1  MAGHTGQGVDFQQIEFAVGI FEEIDAHA AF RTDCLCAANR QFAHQAFFGF
51  GQIFRRTLIN GVVAVVLGFV VVKLGCGDDV YAGQPFVQD GAGIFAAADK
101 TFGNDFAPEG VSILRKRFSD GLFL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 335>:

## m096.seq

```

1  ATGGCTCGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
101 GCCTGCGCGC CGCGAACC GC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTCG CGGTTGTACT
201 TGGCTTCGTA GTCGTAAAC TCGGTTGCGG GAATGATGTG TATGCCGGGC
251 AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 ACGTTCGGCA ATGATTTCGC CC.TGAAGGT GTTTC AATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 336; ORF 096>:

## m096.pep

```

1  MARHTGQGVDFQQIEFAVGI FEEIDAHA AF RTDCLRAANR QFAHQAFFGF
51  GQIFRRTLIN GVVAVVLGFV VVKLGCGNDV YAGQPFVQD GAGIFAAADK
101 TFGNDFAXEG VSILRKRFSD GLFL*

```

m096/g096 96.0% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGI FEEIDAHA AF RTDCLRAANR QFAHQAFFGF GQIFRRTLIN					
	10	20	30	40	50	60
g096	MAGHTGQGVDFQQIEFAVGI FEEIDAHA AF RTDCLCAANR QFAHQAFFGF GQIFRRTLIN					
	70	80	90	100	110	120
m096.pep	GVVAVVLGFV VVKLGCGNDV YAGQPFVQD GAGIFAAADK TFGNDFAXEG VSILRKRFSD					
	70	80	90	100	110	120
g096	GVVSVVLGFV VVKLGCGDDV YAGQPFVQD GAGIFAAADK TFGNDFAPEG VSILRKRFSD					

```

m096.pep  GLFLX
          |||||
g096      GLFLX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 337>:

## a096.seq

```

1  ATGGCCGGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
101 GCCTGCGCGC CGCGAACC GC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTTG CGGTTGTACT
201 TGGCTTCGTA GTCATAAAC TCGGTCGCGG GGATGATGTG TATGCCGGGC
251 AGGCCTTTGC CGTTCAACAC AGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 CCGTTCGGCA ATGATTTCGC CCT.GAAAGT GTTTC AATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 338; ORF 096.ng>:

## a096.pep

```

1  MAGHTGQGVDFQQIEFAVGI FEEIDAHA AF RTDCLRAANR QFAHQAFFGF

```



51 GQIFRRTLIN GVVAVVLGFV VIKLGRGDDV YAGQAFVQH RAGIFAAADK  
101 PFGNDFAXES VSILRKRFSD GLFL\*

m096/a096 92.7% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGIFEEIDAHAAFR	TDCRLAANRQFAHQAFFGFGQIFRRTLIN				
a096	MAGHTGQGVDFQQIEFAVGIFEEIDAHAAFR	TDCRLAANRQFAHQAFFGFGQIFRRTLIN				
	10	20	30	40	50	60
	70	80	90	100	110	120
m096.pep	GVVAVVLGFVVVKLGCGNDVYAGQPFQD	GAGIFAAADKTFGNDFA	XEGVSILRKRFSD			
a096	GVVAVVLGFVVIKLGCGNDVYAGQAFVQH	RAGIFAAADKPF	GNDFAXESVSILRKRFSD			
	70	80	90	100	110	120
m096.pep	GLFLX					
a096	GLFLX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 339>:

g097.seq

1	ATGGATATTT	CAAAACAAAC	ATTGCTGGAT	AGGGTTTTTA	ACCTGAAGGC
51	AAACGGTACG	ACGGTACGTA	CCGAGTTGAT	GGCGGGTTTG	ACGACCTTTT
101	TGACGATGTG	CTACATCGTT	ATCGTCAATC	CCCTGATTTT	GGGCGAGACC
151	GGAATGGATA	TGGGGGCGGT	ATTCGTCGCT	ACCTGTATCG	CATCCGCCAT
201	CGGCTGTTTT	GTCATGGGT	TTATCGGCAA	CTATCCGATT	GCGCTTGCCC
251	CGGGGATGGG	GCTGAATGCC	TATTTACCT	TTGCCGTCGT	TAAGGGTATG
301	GGCGTGCCCT	GGCAGGTGGC	GTTGGGTGCG	GTGTTCAATT	CCGGTCTGAT
351	TTTCATCCTG	TTCAGCTTTT	TTAAAGTCAG	GGAAATGCTG	GTCAACGCAC
401	TGCCATATGG	TTTGAAAATG	TGCATTGCCG	CCGGTATCGG	TTTGTTTTTG
451	GCACTGATTT	CCCTGAAAGG	CGCAGGCATT	ATCGTTGCCA	ATCCGGCAAC
501	CTTGGTCGGC	TTGGGCGATA	TTCATCAGCC	CAGCGCACTG	TTGGCATTGT
551	TCGGTTTTGT	CATGGTGGTC	GTATTGGGGT	ATTTCGCGGT	TCAAGGCGCA
601	ATCATCATCA	CCATTCTGAC	GATTACCGTC	ATTGCCAGCC	TGATGGGTTT
651	GAACGAGTTT	CACGGCGTGG	TCGGCGAAGT	ACCGGGCATT	GCGCCGACCT
701	TTATGCAGAT	GGATTTTAAA	GGTCTGTTTA	CCGTGAGCAT	GGTCAGCGTG
751	ATTTTCGTCT	TCTTCTTGGT	CGATTTGTTC	GACAGTACCG	GAACGCTGGT
801	CGGCGTATCC	CACCGTGCCG	GACTGCTGGT	GGACGGTAAG	CTGCCCCGCC
851	TGAAACGCGC	ACTGCTTGCA	GACTCTACCG	CCATTGTGGC	AGGTGCGGCT
901	TTGGGTACTT	CTTCAACCAC	GCCTTATGTG	GAAAGCGCGG	CGGGCGTATC
951	GGCAGGCGGA	CGGACCGGCC	TGACGGCGGT	TACCGTCGGC	GTATTGATGC
1001	TGGCGTGTCT	GATGTTCTCC	CCATTGGCGA	AAAGTGTTCC	GGTATTTGCC
1051	ACCGCGCCCG	CACTGCTTTA	TGTCGGCACG	CAGATGCTCC	GCAGTGCGAG
1101	GGACATTGAT	TGGGACGATA	TGACTGAAGC	CGCGCCCGCG	TTCCTGACCA
1151	TTGTCTTCAT	GCCGTTTACC	TATTCGATTG	CAGACGGCAT	CGCCTTCGGC
1201	TTCATCAGCT	ATCCGTGGT	CAAACCTTTG	TGTCGCCGGA	CTGGGACGCT
1251	GCCGCCTATG	GTATGGGTTG	TTGCCGTATT	GTGGGCATTG	AAATTCTGGT
1301	ATTTGGGCTG	A			

This corresponds to the amino acid sequence <SEQ ID 340; ORF 097.ng>:

g097.pep

1	MDISKQTL	LD	RVFNLKANGT	TVRTELMAGL	TTFLTMCYIV	IVNPLILGET
51	GMDMGAVFVA	TCIASAIGCF	VMGFIGNYPI	ALAPGMGLNA	YFTFAVVKGM	
101	GVPWQVALGA	VFISGLIFIL	FSFFKVREML	VNALPMGLKM	SIAAGIGLFL	
151	ALISLKGAGI	IVANPATLVG	LGDIHQPSAL	LALFGFVMV	VLGYFRVQGA	
201	IIITILITIV	IASLMGLNEF	HGVVGEVPGI	APTFMQMDFK	GLFTVSMVSV	
251	IFVFFLVDLF	DSTGTLVGVS	HRAGLLVDGK	LPRLKRALLA	DSTAIVAGAA	
301	LGTSTTPYV	ESAAGVSAGG	RTGLTAVTVG	VLMLACLMPF	PLAKSVPVFA	
351	TAPALLYVGT	QMLRSARDID	WDDMTEAAPA	FLTIVFMPFT	YSIADGIAFG	

401 FISYAVVKLL CRRTGDVPPM VWVAVLWAL KFWYLG\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 341>:

m097.seq

```

1 ATGGACACTT CAAAACAAAC ACTGTTGGAC GGGATTTTTA AGCTGAAGGC
51 AAACGGTACK ACGGTGCGTA CCGAGTTGAT GGCGGGTTTG ACAACTTTTT
101 TGACGATGTG CTACATCGTT ATCGTCAACC CTCyGATTTT GGGCGAGACC
151 GGCATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CGTCTGCCAT
201 CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
251 CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
301 GGCGTGCCCT GGCAGGTTGC GTTGGGTGCG GTGTTTCATCT CCGGTCTGAT
351 TTTTATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
401 TGCCTATGGG TTTGAAAATG TCGATTGCTG CCGGTATCGG TTGTTTTTTG
451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
501 CTTGGTCGGT TTGGGCGATA TTCATCAGCC GTCCGCGTTG TTGGCATTGT
551 TCGGTTTTGC TATGGTGGTC GTATTGGGAC ATTTCCGCGT TCAAGGCGCA
601 ATCATCATCA CCATCTTGAC CATTACCGTC ATTGCCAGCC TGATGGGTTT
651 GAATGAATTT CACGGCATCA TCGGCGAAGT ACCGAGCATT GCGCCGACTT
701 TTATGCAGAT GGATTTTGAA GGCCTGTTTA CCGTCAGCAT GGTCAAGTGTG
751 ATTTTCGTCT TCTTCTTGGT CGATCTATTT GACAGTACCG GAACGCTGGT
801 CGGCATATCC CACCGTGCCG GGCTGCTGGT GGACGGTAAG CTGCCCCGCC
851 TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
901 TTGGGTACTT CTTCCACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
951 GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001 TCGCCTGCCT GATGTTTTCA CCTTTGGCGA AAAGTGTTC CGCTTTTGCC
1051 ACCGCGCCCG CCCTGCTTTA TGTGCGCACG CAGATGCTCC GCAGTGCAG
1101 GGATATTGAT TGGGACGATA TGACGGAAGC CGCACCTGCG TTCCTGACCA
1151 TTGTTTTTCAT GCCGTTTACT TATTCGATTG CAGACGGCAT CGCTTTCGGC
1201 TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCCGCA CCAAAGACGT
1251 TCCGCCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT
1301 ATTTGGGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 342; ORF 097>:

m097.pep

```

1 MDTSKQTL LD GIFKLKANGT TVRTEL MAGL TTFLTMCYIV IVNPXILGET
51 GMDMGAVFVA TCIASAIGCF VMGFVGNYP ALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHOPSAL LALFGFAMVV VLGHFRVQGA
201 IIIITILITV IASLMGLNEF HGIIGEVPSI APTFMQMDFE GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGIS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACL MFS PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDMTEAAPA FLTIVFMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDVP PM VWIVAVLWAL KFWYLG*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 097 shows 96.3% identity over a 436 aa overlap with a predicted ORF (ORF 097.ng) from *N. gonorrhoeae*:

m097/g097

	10	20	30	40	50	60
m097.pep	MDTSKQTL LD GIFKLKANGT TVRTEL MAGL TTFLTMCYIV IVNPXILGET GMDMGAVFVA					
g097	MDISKQTL LDRVFN LKANGT TVRTEL MAGL TTFLTMCYIV IVNPLILGET GMDMGAVFVA					
	10	20	30	40	50	60
m097.pep	TCIASAIGCF VMGFVGNYP ALAPGMGLNAYFTFAVVKGM GVPWQVALGAVFISGLIFIL					
g097	TCIASAIGCF VMGFIGNYP ALAPGMGLNAYFTFAVVKGM GVPWQVALGAVFISGLIFIL					
	70	80	90	100	110	120
m097.pep	TCIASAIGCF VMGFVGNYP ALAPGMGLNAYFTFAVVKGM GVPWQVALGAVFISGLIFIL					
g097	TCIASAIGCF VMGFIGNYP ALAPGMGLNAYFTFAVVKGM GVPWQVALGAVFISGLIFIL					
	130	140	150	160	170	180

m097.pep	FSFFKVVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
g097	FSFFKVVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
	130 140 150 160 170 180
m097.pep	LALFGFAMVVVLGHFRVQGAIITILTITVIASLMGLNEFHGIIGEVPSIAPTTFMQMDFE
g097	LALFGFAMVVVLGHFRVQGAIITILTITVIASLMGLNEFHGVVGEVPGIAPTTFMQMDFK
	190 200 210 220 230 240
m097.pep	GLFTVSMVSVIFVFFLVLDLFDSTGTLVGISHRAGLLVDGKLPRLKRALLADSTAIVAGAA
g097	GLFTVSMVSVIFVFFLVLDLFDSTGTLVGVS HRAGLLVDGKLPRLKRALLADSTAIVAGAA
	250 260 270 280 290 300
m097.pep	LGTSSSTTPYVESAAGVSAGGRTGLTAVTVGVMLACL MFSP LAKSVPAFATAPALLYVGT
g097	LGTSSSTTPYVESAAGVSAGGRTGLTAVTVGVMLACL MFSP LAKSVPVFATAPALLYVGT
	310 320 330 340 350 360
m097.pep	QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVPPM
g097	QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTGDVPPM
	370 380 390 400 410 420
m097.pep	VWIVAVLWALKFWYLGX
g097	VWVAVLWALKFWYLGX
	430

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 343>

a097.seq

1	ATGGACACTT	CAAAACAAAC	ACTGTTGGAC	GGGATTTTTA	AGCTGAAGGC
51	AAACGGTACG	ACGGTGCGTA	CCGAGTTGAT	GGCGGGTTTG	ACAACTTTTT
101	TGACGATGTG	CTACATCGTT	ATCGTCAACC	CTCTGATTTT	GGGCGAGACC
151	GGCATGGATA	TGGGGGCGGT	ATTCGTGCGT	ACCTGTATCG	CGTCTGCCAT
201	CGGCTGTTTT	GTTATGGGTT	TTGTGGGCAA	CTATCCGATT	GCACTCGCAC
251	CGGGGATGGG	GCTGAATGCC	TATTTACCT	TTGCCGTCGT	TAAGGGTATG
301	GGCGTGCCCT	GGCAGGTTGC	GTTGGGTGCG	GTGTTTCATCT	CCGGTCTGAT
351	TTTCATCCTG	TTCAGCTTTT	TTAAAGTCAG	GGAAATGCTG	GTCAACGCAC
401	TGCCTATGGG	TTTGAAAATG	TCGATTGCTG	CCGGTATCGG	TTGTTTTTTG
451	GCACTGATTT	CCCTGAAAGG	CGCAGGCATT	ATCGTTGCCA	ATCCGGCAAC
501	CTTGGTCGGC	TTGGGCGATA	TTCATCAGCC	GTCCGCGTTG	TTGGCACTGT
551	TCGGTTTTGC	CATGGTGGTC	GTATTGGGAC	ATTTCCGCGT	TCAAGGCGCA
601	ATCATCATCA	CCATTTTGAC	GATTACCGTC	ATTGCCAGCC	TGATGGGTTT
651	GAACGAATTT	CACGGCATCA	TCGGCGAAGT	GCCGAGCATT	GCGCCGACTT
701	TTATGCAGAT	GGATTTTAAA	GGGTTGTTTA	CCGTCAGCAT	GGTCAGCGTG
751	ATTTTCGTCT	TTTTCTAGT	CGATCTGTTT	GACAGTACCG	GAACACTGGT
801	CGGTGTATCG	CATCGTGCCG	GACTGCTGGT	GGACGGTAAG	CTGCCCGGCC
851	TGAAACGCGC	ACTGCTTGCA	GACTCTACCG	CTATTGTGGC	AGGTGCGGCT
901	TTGGGTACTT	CTTCAACCAC	GCCTTATGTG	GAAAGTGCGG	CGGGCGTATC
951	GGCAGGCGGG	CGGACAGGTC	TGACGGCGGT	TACCGTCGGC	GTATTGATGC
1001	TCGCCCTGCC	GATGTTTTCA	CCTTTGGCGA	AAAGTGTTCC	CGCTTTTGCC
1051	ACCGCGCCCG	CCCTGCTTTA	TGTCGGCACG	CAGATGCTCC	GCACTGCGAG
1101	GGACATCGAT	TGGGACGATA	TGACGGAAGC	CGCACCCGCA	TTCCTGACCA
1151	TTGCTTTCAT	GCCGTTTACC	TATTCGATTG	CAGACGGCAT	CGCTTTCGGC
1201	TTTCATCAGT	ATGCCGTGGT	TAACTTTTAA	TGCCGCGCGA	CCAAAGACGT
1251	TCCGCTATG	GTATGGATTG	TTGCCGTATT	GTGGGCACTG	AAATTCTGGT

1301 ATTTGGGCTG A

This corresponds to the amino acid sequence &lt;SEQ ID 344; ORF 097.a&gt;:

a097.pep

```
1  MDTSKQTLTD GIFKLGKANGT TVRTELMAGL TTFLTMCYIV IVNPLILGET
51  GMDMGAVFVA TCIAAIGCF VMGFVGNYP ALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHFRVQGA
201 IIITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFK GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VMLACLMS PLAKSVPAFA
351 TAPALLYGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDPPM VWIVAVLWAL KFWYLG*
```

m097/a097 99.3% identity in 436 aa overlap

```
          10      20      30      40      50      60
m097.pep  MDTSKQTLTDGIFKLGKANGTTVRTELMAGLTTFLTMCYIVIVNPNXILGETGMDMGAVFVA
          10      20      30      40      50      60
a097      MDTSKQTLTDGIFKLGKANGTTVRTELMAGLTTFLTMCYIVIVNPLILGETGMDMGAVFVA
          70      80      90     100     110     120
m097.pep  TCIAAIGCFVMGFVGNYPALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
          70      80      90     100     110     120
a097      TCIAAIGCFVMGFVGNYPALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
          130     140     150     160     170     180
m097.pep  FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
          130     140     150     160     170     180
a097      FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL
          190     200     210     220     230     240
m097.pep  LALFGFAMVVVLGHFRVQGAIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFE
          190     200     210     220     230     240
a097      LALFGFAMVVVLGHFRVQGAIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFK
          250     260     270     280     290     300
m097.pep  GLFTVSMVSVIFVFFLVDLFDSTGTLVGISHRAGLLVDGKLPRLKRALLADSTAIVAGAA
          250     260     270     280     290     300
a097      GLFTVSMVSVIFVFFLVDLFDSTGTLVGSHRAGLLVDGKLPRLKRALLADSTAIVAGAA
          310     320     330     340     350     360
m097.pep  LGTSSTTPYVESAAGVSAGGRTGLTAVTVGVMLACLMSPLAKSVPAFATAPALLYGT
          310     320     330     340     350     360
a097      LGTSSTTPYVESAAGVSAGGRTGLTAVTVGVMLACLMSPLAKSVPAFATAPALLYGT
          370     380     390     400     410     420
m097.pep  QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDPPM
          370     380     390     400     410     420
a097      QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDPPM
          430
m097.pep  VWIVAVLWALKFWYLGX
          430
a097      VWIVAVLWALKFWYLGX
          430
```

g098.seq

1	ATGACCGCCG	ACGGTCTCTT	CGTCGCTTTC	AACTTCAATA	CGTTTGCCGT
51	TGTGCGAATA	TTGATACCAG	TACAGCAGGA	TGCTGCCCAG	GCTGGCGATC
101	AGTTTGTCGG	CGATGTCGCG	CGCTTCGCTG	TCGGGATGGC	TTTCGCGTTC
151	GGGATGAACG	CAGCCGAGCA	TGGACACGCC	GGTACGCATC	ACGTCCATCG
201	GATGGGTATG	TGCAGGCAGG	CTTTCCAAAA	CTTTAATCAC	ACGGATAGGC
251	AGGCCGCGCA	TGGATTTTAG	CTTGGTTTTA	TAAGCGGCCA	GCTCGAATTT
301	GTTGGGCAGA	TGGCCGTGAA	TCAGCAAGTG	TGCGACTTCT	TCAAACCTCG
351	ATTTTGTGTC	CAAAATTAGAA	TGTCGTAA		

This corresponds to the amino acid sequence <SEQ ID 346; ORF 098.ng>:

g098.pep

1	MTADGLFVAF	NFNTFAVVRI	LIPVQDDAAQ	AGDQFVGDDA	RFAVGMAFAF
51	GMNAAEHGHA	GTHHVHRMGM	CRQAFQNFNH	TDRQAAHGFE	LGFISGQLEF
101	VGOMAVNOOV	CDEEKLAEFC	QIRMS*		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 347>:

m098.seq

```
1 ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
51 TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
101 AGTTTGTCGG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTT
151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTTTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCGAGTG GGCCTACTTCT TCAAACTCGC
351 ATTTTGTGTC CAAATCAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 348; ORF 098>:

m098.pep

1	MTADGLFVAF	NLNAFAVVRI	LIPVQEDAAE	AGDQFVGDDA	RFTFRMAFTF
51	RMNAAQHGYA	GTHYVHRMGM	CRQAFQNFNH	TDRQAAHGFE	LGFISGQLEF
101	VGQMAVNQOV	GDFFKLAFLC	QIRMS*		

Computer analysis of this amino acid sequence gave the following results:

### Homology with a predicted ORF from *N. gonorrhoeae*

ORF 098 shows 89.6% identity over a 125 aa overlap with a predicted ORF (ORF 098.ng) from *N. gonorrhoeae*:

		10	20	30	40	50	60
m098.pep		MTADGLFVAFNLFNAFAVVRI	LIPVQEDAAEAGDQFVG	DVARFTFRMAFTFRMNAAQHGYA			
			: :	: :	:	:	:   :
g098		MTADGLFVAFNFNTFAVVRI	LIPVQQDAAQAGDQFVG	DVARFAVGMAFAFGMNAAEHGH	A		
		10	20	30	40	50	60
		70	80	90	100	110	120
m098.pep		GTHYVHRMGMCRAQAFQNFNHTDRQA	AHGFE	LGFI	SQGLEFVGQMAVNQQVGDFFKLAFLC		
		:					
g098		GTHHVHRMGMCRAQAFQNFNHTDRQA	AHGFE	LGFI	SQGLEFVGQMAVNQQVCDFFKLAFLC		
		70	80	90	100	110	120
m098.pep		QIRMSX					
g098		QIRMSX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 349>:

a098.seq

```
1  ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
51 TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
```

```
101 AGTTTGTGCG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTT
151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GGCGACTTCT TCAAACCTCGC
351 ATTTTGTGTC CAAATCAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 350; ORF 098.a>:

a098.pep

```
1 MTADGLFVAF NLNAFAVVRI LIPVOEDAAE AGDQFVGDDVA RETFRMAFTF
51 RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
101 VGQMAVNQOV GDFFKLAFLC QIRMS*
```

m098/a098 100.0% identity in 125 aa overlap

```

      10      20      30      40      50      60
m098.pep MTADGLFVAFNLNAFAVVRI LIPVOEDAAEAGDQFVGDDVARFTFRMAFTFRMNAQHGYA
          |||||
a098      MTADGLFVAFNLNAFAVVRI LIPVOEDAAEAGDQFVGDDVARFTFRMAFTFRMNAQHGYA
          10      20      30      40      50      60

      70      80      90     100     110     120
m098.pep GTHYVHRMGMCROAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQOVGDFFKLAFLC
          |||||
a098      GTHYVHRMGMCROAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQOVGDFFKLAFLC
          70      80      90     100     110     120

m098.pep QIRMSX
          |||||
a098      QIRMSX
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 351>:

g099.seq

```
1 ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTGGA
51 GCTGACGGGC AAACGGCAGG CGGGCATTAC TGCCACAGAC ATCGTGTGTTG
101 CACTGACCGA ATTCTTGCGT AAAGAGCGCG TGGTCGGGGC GTTTGTGCGAA
151 TTTTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCCATG TTCGCCATCG
251 ACGCGCAAAC TATTGATTAT TTGAAACTGA CCGGACGTGA CGACGCGCAG
301 GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTAT GGGCAGGTGG
351 CTTGAAAACC GCCGTTTATC CGCGCGTTTT GAAATTTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCACGCGCG TTTTGCCACC
451 GCCGATTGCG CGGCGAAAGG GCTGGCGAAG CTTTACGAAG AGCCTTCAGA
501 CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCGTGTA
551 CCAATACTTC CAACCCGCGC AACGTTGTCG CCGCCGCACT GTTGGCACGC
601 AATGCCAACC GCCTCGGCTT GAAACGCAA CTTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGGAATCTA TTTGAAAGAA GCAGGCTTGT
701 TGCCCGAAAT GGAAAACTC GGCTTCGGTA TCGTCGCCTT CGCATGTACC
751 ACCTGTAAAC GCATGAgcgG CGCGCTcgac CCGAAAATCC AACAAGAAAT
801 CATCGACCGC GAttgtacg cCACCgCGT ATTGTCAGGC AACCgCAACT
851 TCGACGCGCG TATCCATCCG TATGCGAAAC AGGCTTTCTT CGCTTCGCCT
901 CCTTTGGTGC TTGCCTACGC ATTGGCAGGT AGCATCCGTT TCGATATTGA
951 AAACGACGTA CTCGGCGTTG CAGACGCGCG CGAAATCCGC CTGAAAGATA
1001 TCTGGCCGAC AGACGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051 CCGCAACAAT TCCGCGACAT TTATATCCCG ATGTCCGACA CCGGCACAGC
1101 GCAAAAAGCA CCAAGCCCGC TGTACGACTG GCGACCGATG TCCACCTACA
1151 TCCGCGGTCC GCCCTATTGG GAAGGCGCAC TGGCAGGGGA ACGTACATTA
1201 AGAGGTATGC GTCCGCGGCG GATTTTGCCC GACAACATCA CCACCGACCA
1251 CATCTCgcca tCCAATGCGA TTTTGGCCGG cagTGcgcga ggtgaATATT
```

```

1301 TGGCGAAAAT GGGTTTGCCT GAAGAagaCT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTtcgt
1451 tggcacgcgT tgaacCAGAA GGCCAAACCA TCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTT ATCATCATTG CCGGTGCGGA
1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC GGCGAAGGGC GTGCGGCTGG
1601 CGGGTGTGGA AGCCATCGCC GCCGAAGGTT TCGAGCGCAT CCACCGCACC
1651 AACCTCATCG GCATGGGCGT CTTGCCGCTG CAATTCAAAC CCGGCACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTGCGCG
1751 AACGCACACC GCGCTGCGGC CTGACCCTCG TGATTACCG TAAAAACGGA
1801 GAAACCGTCG AAGTTCCGGT TACCTGCCGC CCCGATACCG CAGAAGAAGC
1851 ATTGGTATAT GAAGCCGGCG GCGTATTGCA ACGGTTTGCA CAGGACTTTT
1901 TGGAAGGGAA CGCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 352; ORF 099.ng>:

g099.pep

```

1 MLGRASMMRL PDIVGVELTG KRQAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDAQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWAGGLKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAAKGLAK PYEPPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLKRR PWVKSSFAPG SKVAGIYLKE AGLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFASP
301 PLVVAYALAG SIRFDIENDV LGVADGREIR LKDIWPTDEE IDAIVAEYVK
351 PQQFRDIYIP MSDTGTAKA PSPLYDWRPM STYIRPPYW EGALAGERTL
401 RGMRPAILP DNITTDHISP SNAILAGSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501 ETYMNKQPL IIIAGADYGO GSSRDWAAKG VRLAGVEAIA AEGFERIHRT
551 NLIGMGVLP L QFKPGTNRHT LQLDGTETD VVGERTPRCG LTLVIHRKNG
601 ETVEVPVTCR PDTAEALVY EAGGVLRFA QDFLEGNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 353>:

m099.seq

```

1 ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
51 GCTGAACGGC AAACGGCAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
101 CACTGACCGA GTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTGTGCGAA
151 TTCTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGAGT TCGGCGCGAC TGCCGCGATG TTCGTATTG
251 ATGAGCAAAC CATTGATTAT TTGAACTGA CCGGACGCGA CGACGCGCAG
301 GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
351 CTTGAAAACC GCCGTTTATC CTCGCGTTT GAAATTTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCAAGTAACC CGCATGCCCG TTTTGCAGCC
451 GCCGATTGG CGGCGAAAGG GCTGGCGAAG CCTTACGAAG AGCCTTCGGA
501 CGGCCAAATG CCCGACGCT CGGTCATCAT CGCCGCGATT ACCAGTTGCA
551 CCAACACTTC CAACCCGCGC AACGTTGTTG CCGCCGCGCT CTTGGCACGC
601 AATGCCAACC GTCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
651 TGCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCGGGCCTGT
701 TGCCCGAAAT GGA AAAACTC GGCTTCGGTA TCGTCGCCTT CGCTGCACC
751 ACCTGCAACG GCATGAGTGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
801 CATCGACCGC GATTTGTACG CCACCGCGT ATTATCAGGC AACCGCAACT
851 TCGACGGCCG TATCCACCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
901 CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGTATCCGTT TCGATATTGA
951 AAACGACGTA CTCGGCGTTG CAGACGGCAA GGAAATCCGC CTGAAAGACA
1001 TTTGGCTGCG CGATGAAGAA ATCGATGCCG TCGTTGCCGA ATATGTGAAA
1051 CCGCAGCAGT TCCGCGATGT GTATGTACCG ATGTTGACA CCGGCACAGC
1101 GCAAAAAGCA CCCAGTCCGC TGTACGATTG GCGTCCGATG TCCACCTACA
1151 TCCGCCGTCC GCCTTACTGG GAAGGCGCGC TGGCAGGGGA ACGCACATTA
1201 AGAGGTATGC GTCCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA
1251 CCTCTCGCCG TCCAATGCGA TTTTGGCCGT CAGTGCCGCA GGCGAGTATT
1301 TGGCGAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCT ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAAAA ACGAAGACGG CAGCGTGC GC CAAGGCTCGT
1451 TCGCCCGCGT CGAACCCGAA GGCGAAACCA TCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGTGCGGA

```

```

1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601 CCGGCGTAGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACCTTATCG GCATGGGCGT GTTGCCGCTG CAGTTCAAAC CCGACACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTGGTCGGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTCG TGATTCACCG TAAAAACGGC
1801 GAAACCGTTG AAGTTCCCGT TACCTGCTGC CTCGATACTG CAGAAGAAGT
1851 ATTGGTATAT GAAGCCGGCG GCGTGTGCA ACGGTTTGCA CAGGATTTT
1901 TGGAAGGGAA CGCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 354; ORF 099>:

```

m099.pep
1  MLGRASMMRL PDIVGVELNG KRQAGITATD IVLALTEFLR KERVVGAFVE
51  FFGEGARSL S IGDRATISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAAKGLAK PYEEPSDGQM PDGSVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLKRK PWVKSSFAPG SKVAEIYLKE AGLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQKEIIDR DLYATAVL SG NRNFDGRIHP YAKQAFASP
301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPADEE IDAVVAEYVK
351 PQQFRDVYVP MFDTGTAQKA PSPLYDWRPM STYIRRPYEW EGALAGERTL
401 RGMRLAILP DNITTDHLSP SNAILAVSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVKNEDGSVR QGSFARVEPE GETMRMWEAI
501 ETYMNKQPL IIIAGADYGO GSSRDWAAKG VRLAGVEAIV AEGFERIHRT
551 NLIGMGVLPL QFKPDTNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
601 ETVEVPVTC LDTAEEVLVY EAGGVLQRF QDFLEGNA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 099 shows 96.2% identity over a 639 aa overlap with a predicted ORF (ORF 099.ng) from *N. gonorrhoeae*:

```

m099/g099
10      20      30      40      50      60
m099.pep  MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSL
g099      MLGRASMMRLPDIVGVELTGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSL
10      20      30      40      50      60

70      80      90      100     110     120
m099.pep  IGDRATISNMTPEFGATAAMFAIDEQTIDYKLKLTGRDDAQVKLVETYAKTAGLWADALK
g099      IGDRATISNMTPEFGATAAMFAIDAQTIDYKLKLTGRDDAQVKLVETYAKTAGLWAGGLK
70      80      90      100     110     120

130     140     150     160     170     180
m099.pep  AVYPRVLKFDLSSVTRNMAGPSNP HARFATADLAAKGLAKPYEEPSDGQMPDGSVIIAAI
g099      AVYPRVLKFDLSSVTRNMAGPSNP HARFATADLAAKGLAKPYEEPSDGQMPDGA VIIAAI
130     140     150     160     170     180

190     200     210     220     230     240
m099.pep  TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAEIYLKEAGLLPEMEKL
g099      TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAGIYLKEAGLLPEMEKL
190     200     210     220     230     240

250     260     270     280     290     300
m099.pep  GFGIVAFAC TTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFASP
g099      GFGIVAFAC TTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFASP
250     260     270     280     290     300

310     320     330     340     350     360

```



m099.pep	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIIWPADEEIDAVVAEYVKPQQFRDVVYP					
	:           :           :           :					
g099	PLVVAYALAGSIRFDIENDVLGVADGREIRLKDIIWPTDDEIDAIVA EYVKPQQFRDIYIP					
	310	320	330	340	350	360
	370	380	390	400	410	420
m099.pep	MFDTGTAQKAPSPLYDWRPMSTYIRRPPYWEGALAGERTLRGMRPLAILPDNITTDHLS P					
	:           :           :           :					
g099	MSDTGTAQKAPSPLYDWRPMSTYIRRPPYWEGALAGERTLRGMRPPAILPDNITTDHISP					
	370	380	390	400	410	420
	430	440	450	460	470	480
m099.pep	SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVK NEDGSVR					
	:           :           :           :					
g099	SNAILAGSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVR NEDGSVR					
	430	440	450	460	470	480
	490	500	510	520	530	540
m099.pep	QGSGARVEPEGETMRMWEAIETYMN RKQPLIIIAGADYGQGSSRDWA AKGVRLAGVEAIV					
	:     :           :           :           :					
g099	QGSLARVEPEGQTRMWEAIETYMN RKQPLIIIAGADYGQGSSRDWA AKGVRLAGVEAIA					
	490	500	510	520	530	540
	550	560	570	580	590	600
m099.pep	AEGFERIHRTNLIGMGVLP LQFKPDTNRHTLQLDGTET YDVVGERTPRCDLT LV IHRKNG					
	:           :           :           :					
g099	AEGFERIHRTNLIGMGVLP LQFKPDTNRHTLQLDGTET YDVVGERTPRC GLT LV IHRKNG					
	550	560	570	580	590	600
	610	620	630	640		
m099.pep	ETVEVPVTCCLDTAEEVL VYEAGGVLQRFAQDFLEGNAA X					
	:           :           :           :					
g099	ETVEVPVTCRPDTAE EALVYEAGGVLQRFAQDFLEGNAA X					
	610	620	630	640		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 355>:

**a099.seq**

1	ATGCTGGGAC	GCGCGTCCAT	GATGCGCCTG	CCCGATATTG	TCGGCGTTGA
51	GCTGAACGGC	AAACGGAAGG	CGGGCATTAC	GGCGACGGAT	ATTGTGTTGG
101	CACTGACCGA	GTTTCTGCGC	AAAGAACGCG	TGGTCGGGGC	GTTTGTCTGA
151	TTCTTCGGCG	AGGGCGCGAG	AAGCCTGTCT	ATCGGCGACC	GCGCGACCAT
201	TTCCAACATG	ACGCGGAGT	TCGGCGCGAC	TGCCGCGATG	TTCGCTATTG
251	ATGAGCAAAC	CATTGATTAT	TTGAAACTGA	CCGGACGCGA	CGACGCGCAG
301	GTGAAATTGG	TGGAAACCTA	CGCCAAAACC	GCAGGCTTGT	GGGCAGATGC
351	CTTGAAAACC	GCCGTTTATC	CGCGCGTTTT	GAAATTTGAT	TTGAGCAGCG
401	TAA CGCGCAA	TATGGCAGC	CCGAGCAACC	CGCACGCGCG	TTTTGCGCAC
451	GCCGATTTGG	CCGGCAAAGG	CTTGGCTAAA	CCTTACGAAG	AGCCTTCAGA
501	CGGCCAAATG	CCTGACGGTG	CAGTGATTAT	TGCCGCGATT	ACTTCCTGTA
551	CCAATACTTC	CAATCCGCGC	AACGTTGTCT	CCGCCGCGCT	GTTGGCACGC
601	AATGCCAACC	GCCTCGGCTT	GCAACGCAAA	CCTTGGGTGA	AATCTTCGTT
651	TGCCCCGGGT	TCAAAAGTAG	CCGAATCTA	TTTGAAGTAA	CGAGATCTGC
701	TGCCCCGAAAT	GAAAAAATC	GGCTTCGGTA	TCGTTGCCTT	CGCATGTACC
751	ACCTGTAACG	GCATGAGCGG	CGCGCTGGAT	CCGAAAATCC	AGAAAGAAAT
801	CATCGACCGC	GATTTGTACG	CCACCGCCGT	ATTGTCAGGC	AACCGCAACT
851	TTGACGCGCG	TATCCATCCG	TATGCGAAAC	AGGCTTTCCCT	CGCTTCGCCT
901	CCGTTGGTCT	TTGCCATACG	CTGTGGCAGC	AGCATCCGTT	TCGATATTGA
951	AAACGACGTA	CTCGGCGTTG	CAGACGGCAA	AGAAATCCGC	CTGAAAGACA
1001	TTTGGCCTAC	CGATGAAGAA	ATCGATGCCA	TCGTTGCCGA	ATATGTGAAA
1051	CCGCAGCAAT	TTCGCGACGT	TTATATCCCG	ATGTTTCGACA	CCGGCACAGC
1101	GCAAAAAGCA	CCAAGCCCGC	TGTACGACTG	GCGTCCAATG	TCTACCTATA
1151	TCCGCGGCCC	ACCTTACTGG	GGAAGGCGAC	TGGCAGGGGA	ACGCACATTA
1201	AGCGGTATGC	GTCCGCTGGC	GATTTTGCCC	GACAACATCA	CCACCGACCC

```
1251 TCTCTCGCCA TCCAATGCGA TTTTGGCAAG CAGTGCCGCA GGCGAATATT
1301 TGGCAAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGTGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTTTCGC
1451 TGGCACGCGT TGAACCCGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGCGCGGA
1551 CTACGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601 CCGGCGTGGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACTTGATCG GTATGGGCGT GTTGCCGCTG CAGTTCAAAC CGGGTACCAA
1701 CCGCCACACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTGCGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTTG TGATTACCG TAAAAACGGC
1801 GAGACCGTCG AAGTCCCCAT TACCTGCCGC CTCGATACCG CAGAAGAAGT
1851 GTTGGTATAT GAAGCCGGTG GCGTATTGCA ACGGTTTGCA CAGGATTTTT
1901 TGAAGGGAA CGCGCTTAG
```

This corresponds to the amino acid sequence <SEQ ID 356; ORF 099.a>:

a099.pep

```
1 MLGRASMMRL PDIVGVELNG KRKAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAGKGLAK PYEEPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLQRK PWVKSSFAPG SKVAEIYLKE ADLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFLAS
301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPTDEE IDAIVAEYVK
351 PQQFRDVYIP MFDTGTAQKA PSPLYDWRPM STYIRPPYEW EGALAGERTL
401 SGMRLAILP DNITTDHLSF SNAILASSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLENE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501 ETYMRKQPL IIIAGADYGO GSSRDWAAKG VRLAGVEAIV AEGFERIHRT
551 NLIGMGLPL QFKPGTNRHT LQDGTETYD VVGERTPRCD LTLVIHRKNG
601 ETVEVPITCR LDTAEVLVY EAGGVLQRFQ QDFLEGNA*
```

m099/a099 97.5% identity in 639 aa overlap

```
10 20 30 40 50 60
m099.pep MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 MLGRASMMRLPDIVGVELNGKRKAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
10 20 30 40 50 60
70 80 90 100 110 120
m099.pep IGDRATISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT
a099 IGDRATISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT
70 80 90 100 110 120
130 140 150 160 170 180
m099.pep AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDGQMPDGSVIIAAI
a099 AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAGKGLAKPYEEPSDGQMPDGAVIIAAI
130 140 150 160 170 180
190 200 210 220 230 240
m099.pep TSCTNTSNPRNVVAAALLARNANRLGLKRKPWKSSFAPGSKVAEIYLKEAGLLPEMEKL
a099 TSCTNTSNPRNVVAAALLARNANRLGLQRPWKSSFAPGSKVAEIYLKEADLLPEMEKL
190 200 210 220 230 240
250 260 270 280 290 300
m099.pep GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFLAS
a099 GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFLAS
250 260 270 280 290 300
310 320 330 340 350 360
m099.pep PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPADEEIDAIVAEYVKPQQFRDVYVP
```

a099	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPTDEEIDAIVA EYVKPQQFRDVYIP	310	320	330	340	350	360
m099.pep	MFDTGTAQKAPSPLYDWRPMS TYIRRPYEWEGALAGERTLRGMRPLAILPDNITTDHLSP	370	380	390	400	410	420
a099	MFDTGTAQKAPSPLYDWRPMS TYIRRPYEWEGALAGERTLSGMRPLAILPDNITTDHLSP	370	380	390	400	410	420
m099.pep	SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNEDGSVR	430	440	450	460	470	480
a099	SNAILASSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNEDGSVR	430	440	450	460	470	480
m099.pep	QGSFARVEPEGETMRMWEA IETYMNRKQPLII IAGADYQGSSRDWAAKGVRLAGVEAIV	490	500	510	520	530	540
a099	QGSLARVEPEGQTRMWEA IETYMNRKQPLII IAGADYQGSSRDWAAKGVRLAGVEAIV	490	500	510	520	530	540
m099.pep	AEGFERIHRTNLIGMGVLP LQFKPDTNRHTLQLDGTET YDVVGERTPRCDLTLVIHRKNG	550	560	570	580	590	600
a099	AEGFERIHRTNLIGMGVLP LQFKPGTNRHTLQLDGTET YDVVGERTPRCDLTLVIHRKNG	550	560	570	580	590	600
m099.pep	ETVEVPVTCCLDTAE EVLVYEAGGV LQRFAQDFLEGNAAX	610	620	630	640		
a099	ETVEVPITCRLDTA E EVLVYEAGGV LQRFAQDFLEGNAAX	610	620	630	640		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 357>:

```

g102.seq
1      AtgtCCGCCA AAactccgtc gctcttcgcg gccgcgatga Ttatccgcgg
51     gaaggttatc ggcgcAGgta ttttcccaa ccccaccgcg aactgggggg
101    acgggttaat aggcctcgctg ttgtgtgctg tgtacacctc gtttccattc
151    tcctccggcg ccctcatgat ttgtgaagtc aacaccataa accCCcggagg
201    ggcaAGTttt gacaccATGg tCAAagacct gctcgGACGc ggctggaaac
251    tcatcaacgg catcgccgtc gctttggTCc tatacgctc gacctacgcg
301    tacattttag tcggcgggtga cctGACCGCC AAAGGCActg GCAGCGAGT
351    AGGCGGCAAA ATTTGctca CCGTCGGACA actcgttctt CTCGGCATCC
401    TCGCCTTTTG CGTATGGGCA TCCGCGACGT TGCTCGACCG CTTTACCgcg
451    GTCCTCATCG GCGGATGGT ATTAACCTTT ATTTGGGCAA CCGGCGGCCT
501    GGTTCGCCAT GCGAAACCGT CCGTCTCTT CGACACCCAA GCCCCTCGCG
551    GCACCGccta CTGATTtTAC GCGCGCACCG CCTGCCCTCT TGCCTCGCT
601    TCCTTCGGCT TCCACGGCAA CGTTTCCAGC CTGCTCAAAT ACTTTAAAGG
651    CGACGcgCc aaagtGcGA aATCcatctg gGcaggtaCa ttggtTGCCt
701    ttggaattta cgtccTCTgg caaacccgca tCaaagcaa cctTGCcgcc
751    aacgagttcg cCCCcgtgat tgcgcgcgag aggcaactCT CCGTCTGaa
801    tgaaaacctG tcctaaattcg cccaaccgcg cgatatgat aAaatattgt
851    ccctatttcc ctacatggca atcgccacct cctttttagg cgTAACctta
901    ggccgtgttg acaacattcg cgcatacttc aaatggaacg acagatgtgc
951    cggggcggggc accaaaaccg tcgcgctgaa ctctcgcg CCCTgtattt
1001   cctggctgct cctccccacc ggcttcttta ccgccattgg tgcgctccgg
1051   ctggcgccaa ccgctctggga ccaagGcatc attcccgcca tgcgtctcta
1101   cgtttccccc caaaaaattG gcGcaggcaa gacttataAa gtttaCGCG
1151   gcttggtggc gatgttagtc ttcccttttc gcatcgccaa catcgccgca
1201   CAGGTATTGA GccAAatgGa ACTcgtCCCC GTATTAAAG GATAA

```

This corresponds to the amino acid sequence <SEQ ID 358; ORF 102.ng>:

g102.pep

```
1  MSAKTPSLFG GAMIAGKVI GAGMFPNPTA NLGDGLIGSL IVLLYTWFPF
51  SSGALMILEV NTHNPRGASF DTMVKDLLGR GWNIINGIAV ALVLYGSTYA
101 YILVGGDLTA KLGSAVGGK ISLTVGQLVF FGILAFVCWA SARLVDRFTG
151 VLIGGMVLTF IWATGGLVAD AKPSVLFDTO APVGTGYWIY AATALPVCLA
201 SFGFHGNVSS LLKYFKGDAP KVAKSIWAGT LVALVIYVLW QTAIQSNLPR
251 NEFAPVIAAE RQLSVLNETL SKFAQTGDMD KILSLFPYMA IATSFGLVTL
301 GLFDNIADIF KWNDSMSGRG TKTVALNFLP PLISWLLPT GFFTAIGASG
351 LAATVWDQGI IPAMLLYVSP QKIGAGKTYK VYGGLWMLV FLFGIANIAA
401 QVLSQMELVV VFVK*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 359>:

```
m102.seq
1  ATGCCCAACA AAACCCCTTC ACTGTCGGGC GGC GCGATGA TTATCGCCGG
51  CACGGTCATC GCGCAGGCA TGCTCGCCAA CCCGACCGCC ACATCCGGCG
101 TATGGTTTAC CGGCTCGCTG GCCGTGTTGC TGTACACCTG GTTTTCTATG
151 CTTTCCAGCG GCCTGATGAT TTTGGAAGTC AACACCCATT ATCCGCACGG
201 CGCAAGTTTC GACACGATGG TCAAAGACCT GCTCGGACGC GGCTGGAACA
251 TCATCAACGG CATCGCCGTC GCCTTCGTTT TATACCTGCT TACTACGCT
301 TATATCTTCG TCGGCGGCGA CCTGACCGCC AAAGGCTTAG GCAGCGCGGC
351 AGCGCGCGAC GTTTCACCTCA CCGTCGGACA ACTCGTCTTC TTCGGCATCC
401 TCGCCTTTTG CGTATGGGCA TCCGCACGCT TGGTCGACCG CTTACCGGCG
451 GTCCTTATCG GCGGCATGGT ATTGACCTTT ATTTGGGCGG CCGCGGGGCT
501 GATTGCCGAT GCCAAGCCGT CCGTCCTCTT CGATACCCAA GCGCCGCGCG
551 GCACAAACTA CTGGATTAC GCGGCCACCG CCCTGCCCGT CTGCTCGCT
601 TCCTTCGGCT TCCACGGCAA CGTCTCCAGC CTGCTCAAAT ACTTTAAAGG
651 CGACGCGCCC AAAGTGCTA AATCCATCTG GACGGGCACA CTGATTGCGC
701 TGGTAATTTA CGTCTCTG GCAAACGCCA TCCAAGGCAA CCGCGCGCGC
751 AACGAGTTCC CCCCCTCAT CGCGCGCGAA GGGCAAGTCT CCGTCTCAT
801 CGAAACCTG TCCAAATTCG CCCAAACCGG CAATATGGAC AAAATATTGT
851 CCCTGTTTTT CTATATGGCG ATCGCCACCT CGTTTTTAGG CGTAACGCTC
901 GGACTCTTCG ACTACATCGC CGACATCTTC AAATGGAACG ACAGCATCTC
951 CGGCGCGACC AAAACCGCCG CGCTGACCTT CTGCGCGCCC CTGATTTCCT
1001 GCCTGCTCTT CCCACCGGCG TTCGTTACCG CCATCGGCTA CGTCGGCCTG
1051 GCGGCAACCG TCTGGACAGG CATCATCCCC GCCATGCTGC TCTACCGTTC
1101 GCGCAAAAAA TTCGGCGCAG GCAAAACCTA TAAAGTTAC GCGCGCTTGT
1151 GGCTGATGGT TTGGGTCTTC CTTTTCGGCA TCGTCAACAT CGCCGCACAG
1201 GTATTGAGCC AAATGGAACG CGTCCCGGTA TTAAAGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 360; ORF 102>:

```
m102.pep..
1  MPNKTPSLFG GAMIAGTVI GAGMLANPTA TSGVWFTGSL AVLLYTWFSM
51  LSSGLMILEV NTHYPHGASF DTMVKDLLGR GWNIINGIAV AFVLYLLTYA
101 YIFVGGDLTA KLGSAAGGD VSLTVGQLVF FGILAFVCWA SARLVDRFTG
151 VLIGGMVLTF IWAAGGLIAD AKPSVLFDTO APAGTNYWIY AATALPVCLA
201 SFGFHGNVSS LLKYFKGDAP KVAKSIWGT LIALVIYVLW QTAIQGNLPR
251 NEFAPVIAAE QVSVLIETL SKFAQTGNMD KILSLFSYMA IATSFGLVTL
301 GLFDYIADIF KWNDSISGRT KTAALTFLPP LISCLLFTG FVTAIGYVGL
351 AATVWTGIIP AMLLYRSRKK FGAGKTYKVY GGLWLMVWVF LFGIVNIAAQ
401 VLSQMELVV VFVK*
```

m102/g102 86.0% identity in 415 aa overlap

```
m102.pep      10      20      30      40      50      60
|  |||||  |||||  |||||  |||||  |||||  |||||
g102          10      20      30      40      50      60
MSAKTPSLFGGAMIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV

m102.pep      70      80      90     100     110     120
|  |||||  |||||  |||||  |||||  |||||  |||||
g102          70      80      90     100     110     120
NTHYPHGASFDTMVKDLLGRGWNIINGIAVAFVLYLLTYAYIFVGGDLTAKGLSAAGGD

m102.pep      130     140     150     160     170     180
|  |||||  |||||  |||||  |||||  |||||  |||||
g102          130     140     150     160     170     180
ISLTVGQLVFFGILAFVCWASARLVDRFTGVLIGGMVLTFIWAAGGLIADAKPSVLFDTO

m102.pep      190     200     210     220     230     240
|  |||||  |||||  |||||  |||||  |||||  |||||
g102          190     200     210     220     230     240
APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWGTGTLIALVIYVLW
```

[illegible]

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 361>:

**a102.seq**

1	ATGCCACCA	AAACCCCTT	ACTGTTCCGG	GGCGGATGA	TTATCGCCG
51	CACGNTCAT	GGCGCAGGT	TGCTCGCCAA	CCCGACCGCC	ACATCCGGCG
101	TATGGTTTAC	CGGCTCGCTG	GCCGTGATTG	TGTACACCTG	GTTTTCCATG
151	CTCTCCAGCG	GCCTGATGAT	TTTGGAAAGT	AACACCCACT	ACCCCCACGG
201	CGCGANCTT	GACACCATGG	TTAAAGACCT	GCTCGGACGG	AGCTGGAACA
251	TCATCAACGG	CATCGCCGTC	GCCTTCGTTT	TATACCTGCT	TACTTACGCT
301	TATATCTTCG	TCGGCGGCGA	CCTGACCGCC	AAAGGCTTAG	GCAGCGCGGC
351	AGGCGGCAAT	GTTTCACTCA	CCGTCCGACA	ACTCGTCTTC	TTCGGCATTC
401	TCGCTTTTGT	CGTATGGGCA	TCCGCACGCT	TGGTCGACCG	ATTCACCGAC
451	GTCCTCATCG	CGCGCATGGT	ATTAACCTTT	ATTTGGGCAA	CCGGCGGCCCT
501	GATTGCCGAT	GCCAACTGTC	CCGTCTCTT	CGACACCCAA	GCCCCTACCG
551	GCACCAACTA	CTGGATTTAT	GTGCGCACCG	CCCTGCCCGT	CTGCCTTGCG
601	TCATTCCGTT	TCCACGGCAA	CGTCTCCAGC	CTGCTCAAAT	ACTTTAAAGG
651	CGACGCGCCC	AAAGTGGCTA	AATCCATCTG	GACGGGCACA	CTGATTGCGC
701	TGGTAATTTA	CGTCTCTGG	CAACCCGCCA	TCCAANGCAA	CCGTCCGCGC
751	AACGAGTTTC	CCCCCGTGAT	TGCCGCCGAA	GGGCAAGTCT	CGTCNTGAT
801	TGAAACCCTG	TCCAAATTCT	CCCAAACCGG	CAATATGGAC	AAAATATTGT
851	CCCTGTTTTT	CTATATGGCG	ATCGCCACCT	CGTTTTTAGG	CGTAACGCTC
901	GGACTCTTCG	ACTACATCGC	CGACATCTTC	AAATGGAACG	ACAGCGTGTC
951	CGGCGGCACC	AAAACCGCGG	GCCTGACCTT	CTCGCGCCT	NTAATTTCTT
1001	GCCTGCCTTT	CCCCACCGGC	TTTGTTACCG	CCATCGGNTA	CGTCGGCCTG
1051	GCGGCAACCG	TCTGGACAGG	CATCATCCCC	GCCATGCTGC	TNTACCGTTC
1101	GCGCAAAAAA	TTCCGGCGCAG	GCAAAACCTA	TAAAGTTTAC	GCGGCGTTGT
1151	GGCTGATGGT	TTGGGTCTTC	CTTTTCGGCA	TCNTCAACAT	CGCCGCACAN
1201	GTATTGAGCC	AAATGGAAC	CGTCCCCGTA	TTTAAAGGAT	AA
1202					

This corresponds to the amino acid sequence <SEQ ID 362; ORF 102.a>:

**a102.pep**

1	MPTKTPSLEF	GAMIIAGTXI	GAGMLANPTA	TSGVWFTGSL	AVLLYTWFMS
51	LSSGLMLEV	NTHYPHGAXF	DTMVKDLLGR	SWNIINGIAV	AEVLYLLTYA
101	YIFVGGDLTA	KGLGSAAGGN	VSLTVGQLVF	FGIALFCVWA	SARLVDRFTS
151	VLIGGMVLTF	ILWATGLLIAD	AKLPVLDFDQ	APTGTNYWII	VATALPVCLA
201	SFGFHGNVSS	LLKYFKGDAP	KVAKSIWTGT	LIALVIYVLW	QTAIQXNLPR
251	NEFAPVIAAE	QGVSVXIETL	SKFAQTGNMD	KILSLFSYMA	IATSFGLGVTL
301	GLFDYIADIF	KWNDSVSGRT	KTAALTFLPP	XISCLLFPTG	FVTAIGYVGL
351	AATVWGTIIP	AMLLYRSRKK	FGAGKTYKVY	GGLWLMVWVF	LFGIXNIAAX
401	VLSOMELVPV	FKG*			

**m102 / a102** 95.9% identity in 413 aa overlap

m102.pep	10	20	30	40	50	60
	MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
a102	MPTKTPSLFGGAMIIAGTXIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
	10	20	30	40	50	60
m102.pep	70	80	90	100	110	120
	NTHYPHGASFDTMVKDLLGRGWNIIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGD					
a102	NTHYPHGAXFDTMVKDLLGRSWNIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGN					
	70	80	90	100	110	120
m102.pep	130	140	150	160	170	180
	VSLTVGQLVFFGILAFCVWASARLVDRFTGVLIGGMVLTFIWAAGGLIADAKPSVLFDTQ					
a102	VSLTVGQLVFFGILAFCVWASARLVDRFTSVLIGGMVLTFIWATGGLIADAKLPVLFDTQ					
	130	140	150	160	170	180
m102.pep	190	200	210	220	230	240
	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTIALVIYVLW					
a102	APTGTNYWIYVATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTIALVIYVLW					
	190	200	210	220	230	240
m102.pep	250	260	270	280	290	300
	QTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
a102	QTAIQXNLPRNEFAPVIAAEGQVSVXIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
	250	260	270	280	290	300
m102.pep	310	320	330	340	350	360
	GLFDYIADIFKWNDISISGRTKTAALTFLPPLISCLLFPTGFVTAIGYVGLAATVWTGIIP					
a102	GLFDYIADIFKWNDISVSGRTKTAALTFLPPXISCLLFPTGFVTAIGYVGLAATVWTGIIP					
	310	320	330	340	350	360
m102.pep	370	380	390	400	410	
	AMLLYRSRKKFGAGKTYKVYGGGLWMVWVFLFGIVNIAAQVLSQMEI LVPVFKGX					
a102	AMLLYRSRKKFGAGKTYKVYGGGLWMVWVFLFGIXNIAAXVLSQMEI LVPVFKGX					
	370	380	390	400	410	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 363>:

g105.seq

```

1  Atgtccgcag aaaCATACAc acAAAtcggc tGGgtaggct taggGcaaAt
51  gGgtctgcct atgGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101 TATACAACCG CTCGCCGAC AAAACTGCCC CCATCTCcgC CAAAGGAGCA
151 AAAGTTTACG GCagcACCGC CGAACTCGTC CGCGCTGCC CCGTCATTTT
201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251 GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301 ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
351 TGCCGAAGCA CCCGTTTCCG GATCGGTCGG ACCCGCCACC AACGGCACAC
401 TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
451 ATATTTTCCC TTGTCGGCAA AAAAACCTTC CATTTCCGCG ATGTCGGCAA
501 AGGCTCGGGC GCGAAACTCG TCTTGAATC GCTCTTAGGC ATTTTCGGCG
551 AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
601 GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCTATGTT
651 TCAAACAAAA AAATCACTAT GGGCAAACCG TGAGTTCCCC CTGCCTTTG
701 CACTCAAACA CGTTTCCAAA GAcctTAACC TCGcegtCAA AGAGCTTGAA
751 CAGGCAGGCA ACACCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801 CAAAGCAGTT GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
851 TGAAATTGGC AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 364; ORF 105.ng>:

```
g105.pep
  1  MSAETYTQIG WVGLGQMGLP MVTRLDDGGI EVGVYNRSPD KTAPISAKGA
 51  KVGYSTAELV RACPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
101  TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151  IFSLVGKKTF HFGDVGKGSG AKLVLSNLLG IFGEAYSEAM LMARQFGIDT
201  DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251  QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 365>:

```
m105.seq
  1  ATGTCCGCAA ACGAATACGC ACAAATCGGC TGGaTAGGCT TAGGGCAAAT
 51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101  TATACAACCG CTCGCCCAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
151  AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201  CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251  GCGACGGATT GGCCGGCAAm ATCATCGTCA ACATGAGCAC CATCTCCCCG
301  ACCGAAAaGC TCGCCGTCAA AGCACTTGTC GAAGCGCAGm GaCAGTTTGC
351  CGAAGCACCC GTTTCGGGAT CGGTCGGGCC CGCCACCAAC GGCACGCTGC
401  TGATTCTGTT CGGCGGCAGC GAAcCGtTTT AAACCCGCTG CAAAAAATAT
451  TTTCCCTCGT CGGCAAAAAA ACCTTCCATT TCGGCGATGT CGGCAAAGGT
501  TCGGGCGCGA AACTCGTCTT GAACGCTC TTTGGGCATT TCGGCGAaCG
551  TACAGCGAAs GmTgCTGATG GCGCGGCAGT TCGGCATCGA TACCGACACC
601  ATCGTCGAAG CCATCGGsGA CTCGGCAATG GACTCGCCCA TGTTCCAAAC
651  CAAAAAATCC CTGTGGGCAA ACCGCGAATT CCCGmCCGmC TTCGCCCTCA
701  AACACGCCTC CAAAGACCTC AACCTCGCCG TCAAAGAGCT TGAACAGGCA
751  GGCAACACCC TGCCCCCGGT CGAAACCGTT GCTGCCAGCT ACCGCAAAGC
801  AGTCGAAGCC GGCTACGGGA CACAGGACGT TTCCGGCGTT TACCTGAAAC
851  TGGCAGAACA CTGA
```

This corresponds to the amino acid sequence <SEQ ID 366; ORF 105>:

```
m105.pep
  1  MSANEYAQIG WIGLGQMGLP MVTRLDDGGI EVGVYNRSPD KTAPISAKGA
 51  KVGYNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGX IIVNMSTISP
101  TEKLAVKALV EAQRQFAEAP VSGSVGPATN GTLLILFGGS EPFXTRCKKY
151  FPSSAKKPSI SAMSAKVRAR NSSXTRSWAF SANVQRXXLM ARQFGIDTDT
201  IVEAIGDSAM DSPMFQTKKS LWANREFPXX FALKHASKDL NLAVKELEQA
251  GNTLPAVETV AASYRKAVEA GYGTQDVSGV YLKLAEH
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 105 shows 79.9% identity over a 289 aa overlap with a predicted ORF (ORF 105.ng) from *N. gonorrhoeae*:

```
m105/g105

      10      20      30      40      50      60
g105.pep  MSAETYTQIGWVGLGQMGLPMVTRLDDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV
          |||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
m105      MSANEYAQIGWIGLGQMGLPMVTRLDDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
          10      20      30      40      50      60

      70      80      90     100     110     120
g105.pep  RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
          | |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
m105      RDYPVIFLMVSDYAAVCDILNGVRDGLAGXIIIVNMSTISPTENLAVKALVEAQR-QFAEA
          70      80      90     100     110

      130     140     150     160     170     180
g105.pep  PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSGAKLVLSNLLG
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
m105      PVSGSVGPATNGTLLILFGGSEPFFXTRCKKYFPSSAKKP-SISAMSAKVRARNSSXTRSW
          120     130     140     150     160     170
```

```
          190      200      210      220      230      240
g105.pep  IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
          |:  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
m105      AFSANVQRXXLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPXXFALKHASK
          180      190      200      210      220      230

          250      260      270      280      289
g105.pep  DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEH
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
m105      DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGTQDVSGVYLKLAEH
          240      250      260      270      280
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 367>:

a105.seq

```
1  ATGTCCGCAA  ACGAATACAC  ACAAATCGGC  TGGATAGGCT  TAGGGCAAAT
51  GGGTCTGCCT  ATGGTAACGC  GGCTCTTGGA  CGGCGGCATC  GAAGTCGGCG
101 TATACAACCG  CTCGCCCAC  AAAACTGCCC  CCATCTCCGC  CAAAGGCGCA
151 AAAGTTTACG  GCAACACCG  CGAACTCGTC  CGCGACTATC  CCGTCATTTT
201 CCTGATGGTT  TCCGACTATG  CCGCCGTGTG  CGACATCCTG  AACGGAGTCC
251 GCGACGGATT  GGCCGGCAA  ATCATCGTCA  ACATGAGCAC  CATCTCCCCG
301 ACCGAAAACC  TCGCCGTCAA  AGCACTTGTC  GAAGCCGCAG  GCGGACAGTT
351 TGCCGAAGCA  CCCGTTTCCG  GATCGGTCGG  GCCCGCCACC  AACGGCACGC
401 TGCTGATTCT  GTTCGGCGGC  AGCGAAGCCG  TTTTAAACCC  GCTGCAAAAA
451 ATATTTTCCC  TCGTCGGCAA  AAAAACCTTC  CATTTTCGGC  ATGTCGGCAA
501 AGGTTTCGGC  GCGAAACTCG  TCTTGAATC  GCTCTTGGGC  ATTTTCGGCG
551 AAGCGTACAG  CGAAGCGATG  CTGATGGCGC  GGCAGTTCGG  CATCGATACC
601 GACACCATCG  TCGAAGCCAT  CGGCGGCTCG  GCAATGGACT  CGCCCATGTT
651 CCAAACCAAA  AAATCCCTGT  GGGCAAACCG  CGAATTCCCA  CCCGCTTCG
701 CCCTCAAACA  CGCCTCCAAA  GACCTCAACC  TCGCCGTCAA  AGAGCTTGAA
751 CAGGCAGGCA  ACACCTGCC  CGCCGTCGAA  ACCGTTGCTG  CCAGCTACCG
801 CAAAGCAGTC  GAAGCCGGCT  ACGGCGAACA  GGACGTTTCC  GCGGTTTACC
851 TGAATTGGC  AGAACACTGA
```

This corresponds to the amino acid sequence <SEQ ID 368; ORF 105.a>:

a105.pep

```
1  MSANEYQIG  WIGLGQMLP  MVTRLLDGGI  EVGVYNRSPD  KTAPISAKGA
51  KVGNTAELV  RDYPVIFLMV  SDYAAVCDIL  NGVRDGLAGK  IIVNMSTISP
101 TENLAVKALV  EAAGGQFAEA  PVSGSVGPAT  NGTLLILFGG  SEAVLNPLQK
151 IFSLVGKKT  FFGDVGKGS  AKLVNLSLLG  IFGEAYSEAM  LMARQFGIDT
201 DTIVEAIGGS  AMDSPMFQTK  KSLWANREFP  PAFALKHASK  DLNLAVKELE
251 QAGNTLPAVE  TVAASYRKAV  EAGYGEQDVS  GVYLKLAEH*
```

m105/a105 96.5% identity in 289 aa overlap

```
          10      20      30      40      50      60
m105.pep  MSANEYAQIGWIGLGQMLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
          |||||  |||||  |||||  |||||  |||||  |||||
a105      MSANEYTQIGWIGLGQMLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
          10      20      30      40      50      60

          70      80      90      100     110     119
m105.pep  RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAG-QFAEA
          |||||  |||||  |||||  |||||  |||||  |||||
a105      RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
          70      80      90      100     110     120

          120     130     140     150     160     170     179
m105.pep  PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSGAKLVNLSLLG
          |||||  |||||  |||||  |||||  |||||  |||||
a105      PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSGAKLVNLSLLG
          130     140     150     160     170     180

          180     190     200     210     220     230
```



```

m105.pep      IFGDV-QRXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPPXAFALKHASK
|||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a105          IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
              190      200      210      220      230      240

m105.pep      240      250      260      270      280
DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a105          DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
              250      260      270      280      290

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 369>:

g105-1.seq

1	ATGTCGCCAG	AAACATACAC	ACAAATCGGC	TGGGTAGGCT	TAGGGCAAAT
51	GGGTCTGCCT	ATGGTAACGC	GGCTCTTGGA	CGCGCGCATC	GAACTCGGCG
101	TATACAAACG	CTCGCCCGAC	AAAAC TGCC	CCATCTCCGC	CAAGGAGGCA
151	AAAGTTTACG	CGACGACCGC	CGAACTCGTC	CGCGCTGCC	CCGTCAATTT
201	CCTGATGGTT	TCCGACTATG	CCGCGGTGTG	CGACATCCTG	AACGGAGTCC
251	GCGACGGATT	GGCCGGCAAA	ATCATCGTCA	ACATGAGCAC	CATCTCCCCC
301	ACCGAAAAACC	TCGCGCTCAA	AGCACTTGTC	GAAGCGCGAG	CGGGACGAGT
351	TGCCGAAGCA	CCCGTTTCCG	GATCGGTGCG	ACCCGCCACC	AACGGCACAC
401	TGCTGATTCT	GTTCCGCGGC	AGCGAAGCCG	TTTTAAACCC	GCTGCAAAAA
451	ATATTTTCCC	TTGTGCGCAA	AAAAACCTTC	CATTTCGGCG	ATGTCGGCAA
501	AGGCTTGGGC	CGGAAACTCG	TCTTGAACTC	GCTCTTAGGC	ATTTTCGGCG
551	AAGCGTACAG	CGAAGCGATG	CTGATGCGCG	GGCAGTTCGG	CATCGATACC
601	GACACCATCT	TCGAAGCCAT	CGCGGGCTCG	GCAATGGACT	CGCCTATGTT
651	TCAAACAAAA	AAATCACTAT	GGGCAAAACG	TGAGTTCCCC	CTGCTCTTTG
701	CACTCAAACA	CGCTTCCAAA	GACCTTAACC	TCGCCGTCAA	AGAGCTTGAA
751	CAGGACGGCA	CAACCTTGCC	CGCGCTCGAA	ACCGTTGTGT	CGAGCTACCG
801	CAAAGCAGTT	AGACCCGGTG	ACGGCGAACA	GGACGTTTCC	GCGCTTTACC
851	TGAAATTGGC	AGAACACTGA			

This corresponds to the amino acid sequence <SEQ ID 370; ORF 105-1.ng>:

g105-1.pwp

1	MSAETYYTQIG	WVGLGQMGLP	MVTRLDDGGI	EVGVYNRSPD	KTAPISAKGA
51	KVYGGSTAEVL	RACPVIFLMV	<u>SDYA</u> AVCD <u>IL</u>	NGVRDLGLAK	IIVNMSTSP
101	TENLAVLAKV	EAAAGGQFAEA	SPVSGVGPAT	NGTLLILFGG	SEAVLNLPQK
151	IFSLVGKKTF	HFGDVCKGSG	AKLVLSNLLG	IFGEAYSEAM	LMARQFGIDT
201	DTIVEAIGGS	AMDSMPFQTK	KSLWANREFP	PAFALKHASK	DLNLAVKELE
251	QAGNTLLPAVE	TVAASYRKAV	EAGYGEQDVS	GVYLKLAH*	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 371>:

**m105-1.seq**

1	ATGTCGCGAA	ACGAATACGC	ACAAATCGGC	TGGATAGGCT	TAGGGCAAA
51	GGGTCTGCCT	ATGGTAAACG	GGCTCTTGGA	CGCGCGCATC	GAAGTCGGCG
101	TATACAAACG	GTCGCCCCAG	AAACTGCCC	CCATCTCCGC	CAAGGCGCGA
151	AAAGTTTACG	CACAACACGC	CGAACTCGTC	CGCGACTATC	CCGTCATTTT
201	CCTGATGGTT	TCCGACATATG	CCGCGGTGTA	CGACATCCCTG	AACGGAGTCC
251	GCGACGGATT	GGCCGGCAAA	ATCATCTGTA	ACATGAGACAT	CATCTCCCCG
301	ACCGAAAACC	TCGCGCTCAA	AGCACTTGTC	GAAGCCGCAG	GGCGACAGTT
351	TGCCGAAGCA	CCCGTTTCCG	GATCGGTGCG	GCCCGCCACC	AACGGCACGC
401	TGCTGATTCT	GTCTCGCGCG	AGCGAAGCCcG	TTTTAAACCC	GCTGCAAAAA
451	ATATTTTCCC	TGTCGCGCAA	AAAAACCTTC	CATTTCGGCG	ATGTCGGCAA
501	AGGTTCGGGC	GCGAAACTCG	TCCTGAACTC	GCTCTTGGGC	ATTTTCGGCG
551	AAGCGTACAG	CGAAnCGATG	CTGATGGCGC	GGCAGTTCCG	CATCGATACC
601	GACACCATCG	TCGAAGCCAT	CGGSGACTCG	GCAATGGACT	CGCCCCATGTT
651	CCAAACCAAA	AAATCCCTGT	GGGCAAAACG	CGAATTCCCG	CCCGCCTTCG
701	CCCTCAAACA	CGCCTCCAAA	GACCTCAACC	TCGCGTCAA	AGAAGCTTGAA
751	CAGGACGGCA	AGACCTGCC	CGCGTCGAA	ACCGTTGCTG	CCAGCTACCG
801	CAAGCAGTC	CAACCCGGTG	ACGGCGAACA	GGACGTTTCC	GGCGTTTACC
851	TGAAACTGGC	AGAACACTGA			

This corresponds to the amino acid sequence <SEQ ID 372; ORF 105-1>:

m105-1.pap

1 MSANEYAQIG WIGLGQMGLP MVTRLDDGGI EVGVYNRSPD KTAPISAKGA  
51 KVGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP

101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK  
 151 IFSLVGKKTFF HFGDVGKGS G AKLVNLSLLG IFGEAYSEX M LMARQFGIDT  
 201 DTIVEAIGDS AMDSPMFQTK KSLWANREFF PAFALKHASK DLNLAVKELE  
 251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH\*

m105-1/g105-1 96.9% identity in 289 aa overlap

	10	20	30	40	50	60
m105-1.pep	MSANEYAQIGWIGLGMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
g105-1	MSAETYTQIGWVGLGMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAEVLV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m105-1.pep	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA					
g105-1	RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m105-1.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGS GAKLVNLSLLG					
g105-1	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGS GAKLVNLSLLG					
	130	140	150	160	170	180
	190	200	210	220	230	240
m105-1.pep	IFGEAYSEXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFFPAFALKHASK					
g105-1	IFGEAYSEAMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFFPAFALKHASK					
	190	200	210	220	230	240
	250	260	270	280	290	
m105-1.pep	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVS G VYLKLAEHX					
g105-1	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVS G VYLKLAEHX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 373>:

a105-1.seq

1 ATGTCCGCAA ACGAATACAC ACAAATCGGC TGGATAGGCT TAGGGCAAAT  
 51 GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAAGTCGGC  
 101 TATACAACCG CTCGCCCGAC AAAACTGCC CCATCTCCGC CAAAGCGGCA  
 151 AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT  
 201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC  
 251 GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG  
 301 ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT  
 351 TGCCGAAGCA CCCGTTTCCG GATCGGTCGG GCCCGCCACC AACGGCACGC  
 401 TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA  
 451 ATATTTTCCC TCGTCGGCAA AAAAACCTTC CATTTCCGGC ATGTCGGCAA  
 501 AGGTTCGGGC GCGAACTCG TCTTGAATC GCTCTTGGGC ATTTTCGGCG  
 551 AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC  
 601 GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCCATGTT  
 651 CCAAACCAAA AAATCCCTGT GGGCAAACCG CGAATCCCA CCCGCCTTCG  
 701 CCCTCAAACA CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA  
 751 CAGGCAGGCA ACACCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG  
 801 CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTC GCGGTTTACC  
 851 TGAAATTGGC AGAACACTGA

This corresponds to the amino acid sequence <SEQ ID 374; ORF 105-1.a>:

a105-1.pep

1 MSANEYTQIG WIGLGMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA  
 51 KVYGNTELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIIVNMSTISP  
 101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK  
 151 IFSLVGKKTFF HFGDVGKGS G AKLVNLSLLG IFGEAYSEAM LMARQFGIDT  
 201 DTIVEAIGGS AMDSPMFQTK KSLWANREFF PAFALKHASK DLNLAVKELE  
 251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS G VYLKLAEH\*

a105-1/m105-1 99.0% identity in 289 aa overlap

	10	20	30	40	50	60
a105-1.pep	MSANEYTQIGWIGLGMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
	10	20	30	40	50	60

m105-1	MSANEYAQIGWIGLGQMG LPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
	10 20 30 40 50 60
a105-1.pep	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
m105-1	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
	70 80 90 100 110 120
a105-1.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKSGAKLVLSLLG
m105-1	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKSGAKLVLSLLG
	130 140 150 160 170 180
a105-1.pep	IFGEAYSEAMLMARQFGIDTDITVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
m105-1	IFGEAYSEXMLMARQFGIDTDITVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK
	190 200 210 220 230 240
a105-1.pep	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAHX
m105-1	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAHX
	250 260 270 280 290

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 375>:

```

g107.seq
1  ATGGTATTAA CCTTTATTG GGCACCGGC GGCCTGGTTG CCGATGCCAA
51  ACCGTCCGTC CTCTTCGACA CCCAAGCCCC CGTCGGCACC GGCTACTGGA
101 TTTACGCCGC CACCGCCCTG CCCGTCCTGC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGTTT CCAGCCTGCT CAAATACTTT AAAGGCGACG cgcCcaaagt
201 GgCGAaATCc atctggGcag gtacattggT TGCCttggtta atttacgtcc
251 TCTggcaaac cgccatCcaa agcaaccTGC cgcgcaacga gttcgcCCCC
301 gtgattgccg ccgagaggca actCTCCGTC CTgaatgaaa cccTGtccaa
351 attcgcccaa accggcgata tggataAaat attgtcccta tttccctaca
401 tggcaatcgc cacctccttt ttaggcgTAA Ccttaggcct gtttgacaac
451 atcgccggac atcttcaaat ggaacgacag tatgtccggg cggcaccaaa
501 accgtcgcgc tga

```

This corresponds to the amino acid sequence <SEQ ID 376; ORF 107.ng>:

```

g107.pep
1  MVLTFIWATG GLVADAKPSV LFDTOAPVGT GYWIYAATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWAGTLVALV IYVLWQTAIQ SNLPRNEFAP
101 VIAAERQLSV LNETLSKFAQ TGDMDKILSL FPYMAIATSF LGVTLGLFDN
151 IAGHLQMERQ YVRAAPKPSR *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 377>:

```

m107.seq
1  ATGGTATTGA CCTTTATTG GCGCGCCGGC GGGCTGATTG CCGATGCCAA
51  GCCGTCCGTC CTCTTCGATA CCCAAGCCCC CGCCGGCACA AACTACTGGA
101 TTTACGCCGs CACCGCCCTG CCCGTCCTGC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGTTT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGACGG GCACACTGAT TGCGCTGGTA ATTTACGTCC
251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301 GTCATCGCCG CCGAAGGGCA AGTCTCCGTC CTCATCGAAA CCCTGTCCAA
351 ATTCGCCCAA ACCGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA
401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451 ATCGCCCATC TTCAAATGGA ACGACAGCAT CTCCGGgCCG CACCAAAACC
501 GCCGCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 378; ORF 107>:

```

m107.pep..
1  MVLTFIWAAG GLIADAKPSV LFDTOAPAGT NYWIYAXTAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP

```

101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTGLGFDY  
151 IAHLMERQH LRAAPKPPR\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 107 shows 89.4% identity over a 170 aa overlap with a predicted ORF (ORF 107.ng) from *N. gonorrhoeae*:

m107/g107

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTPAQAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
g107	MVLTFIWAAGGLIADAKPSVLFDTPAQAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
g107	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAERQSVLNETLSKFAQ					
	70	80	90	100	110	120
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYMAIATSF LGVTGLGFDYIA-HLMERQHLRAAPKPPR					
g107	TGDMDKILSLFPYMAIATSF LGVTGLGFDNIAGHLMERQYVRAAPKPSR					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 379>:

a107.seq

```

1  ATGGTATTAA CCTTTATTTG GCAACCGGC GGCCTGATTG CCGATGCCAA
51  ACTGCCCCGTC CTCTTCGACA CCCAAGCCCC TACCGGCACC AACTACTGGA
101 TTTATGTCGC CACCGCCCTG CCCGTCTGCC TTGCGTCATT CCGTTTCCAC
151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGGACGG GCACACTGAT TGCGCTGGTA ATTTACGTCC
251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301 GTGATTGCCG CCGAAGGGCA AGTCTCCGTC CTGATTGAAA CCCTGTCCAA
351 ATTCGCCCAA ACCGGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA
401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451 ATCGCCGACA TCTTCAAATG GAACGACAGC GTGTCCGGCC GCACCAAAC
501 CGCCGCGCTG ACCTTCCTGC CGCCTCTAAT TTCCTGCCTG CTCTCCCCA
551 CCGGCTTTGT TACCGCCATC GGCTACGTCG GCCTGGCGGC AACCGTCTGG
601 ACAGGCATCA TCCCCGCCAT GCTGCTCTAC CGTTCGCGCA AAAAATTCGG
651 CGCAGGCAA ACCTATAAAG TTTACGGCGG CTTGTGGCTG ATGGTTTGGG
701 TCTTCCTTTT CGGCATCGTC AACATCGCCG CACAGGTATT GAGCCAAATG
751 GAACTCGTCC CCGTATTATA AGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 380; ORF 107.a>:

a107.pep

```

1  MVLTFIWAATG GLIADAKLPV LFDTPAQAGT NYWIYVATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP
101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTGLGFDY
151 IADIFKWNDS VSGRTKTAAL TELPPLISCL LFPTGFVTAI GYVGLAATVW
201 TGIIPAMLLY RSRKKFGAGK TYKVYGGWL MVVVFLEFIV NIAAQVLSQM
251 ELVPVFKG*

```

m107/a107 94.8% identity in 154 aa overlap

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTPAQAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
a107	MVLTFIWAATGGLIADAKLPVLDTPAQAGTNYWIYVATALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60

	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
a107						
	70	80	90	100	110	120
	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYMAIATSFLGVTLGLFDYIAHLQMERQHRAAPKPPRX					
a107						
	130	140	150	160	170	180
	TGNMDKILSLFSYMAIATSFLGVTLGLFDYIADIFKWNDSVSGRRTKAALTFLPPLISCL					
	190	200	210	220	230	240
a107	LFPTGFVTAIGYVGLAATVWTGIIPAMLLYRSRKKFGAGKTYKVYGGGLWLMVWVFLFGIV					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 381>:

```

g108.seq
1  ATGttgccgg gCTTCAACCG GATATTCAaa cggTTTGCTC CAACACTCGG
51  AACggCGCAT AAAACGCCgc cCTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATCCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGGCG GCAATCTTAC TCGGCGGGTG
201 CGCCGCCGGC GGCAACACAT TCGGCAGCTT AGACGGCGGC ACGGGTATGG
251 GTGGCAGCAT CGTCAAAATG ACGGTAGAAA gccAATGCCG TGCGGAATTG
301 GACAGGCGCA GCGAATGGCG TTTGACCGCG CTGGCGATGA GTGCCGAAAA
351 ACAGGCGGAA TGGGAAAACA AGATTGCGG CTGCGCTACC GAAGAAGCAC
401 CTAACCAGCT GACCGGCAAC GATGTGATGC AGATGCTGAa ccagtccacG
451 CGCaatcagg cacTtgccgc CctgaccgTC AAAacgggTT CcgctgcTT
501 CAaacgcctg tACCGCTaa

```

This corresponds to the amino acid sequence <SEQ ID 382; ORF 108.ng>:

```

g108.pep
1  MLPGFNRIFK RFAPTLGTAH KTPPFALSRT GRLIRSYRHK RRGFNRKGIE
51  MNKTLNILPA AILLGGCAAG GNTFGSLDGG TGMGGSIVKM TVESQCRAEL
101 DRRSEWRLTA LAMSAEKQAE WENKICGCAT EEAPNQLTGN DVMQMLNQST
151 RNQALAAALTV KTVSACFKRL YR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 383>:

```

m108.seq
1  ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTT CAACACTCGG
51  AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGGTG GCAATCTTAC TCGGCGGCTG
201 CGCCGCCGGA GCGGTAACA CATTCGGCAG CTTAGACGGT GGCACAGGCA
251 TGGGCGGCAG CATCGTCAA ATGGCGGTTG GGAGCCAATG CCGTGCGGAA
301 TTGGACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
351 AAAACAGGCG GAGTGGGAAA ACAAGATTG CGCTTGCCTC GCCCAAGAAG
401 CACCCGAACG GATGACCGGC AACGATGTGA TGCAGATGCT GGCTCCGTCC
451 ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
501 CTTCAAACAC CTGTACCGCT AA

```

This corresponds to the amino acid sequence <SEQ ID 384; ORF 108>:

```

m108.pep
1  MLPGFNRIFK RFVPTLGTAH KTPPFALSRT GRLIRFYRHK RRGFNRKGIE
51  MNKTLNILPV AILLGGCAAG GGNTFGSLDG GTMGGSIVK MAVGSQCRAE
101 LDKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPERMTG NDVMQMLAPS
151 TRNQALAAAL AKTVSACFKH LYR*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 108 shows 89.6% identity over a 173 aa overlap with a predicted ORF (ORF 108.ng) from *N. gonorrhoeae*:

m108/g108

	10	20	30	40	50	60
m108.pep	MLPGFNRIKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKLSILPV					
g108	MLPGFNRIKRFAPTLGTAKHTPPFALSRTGRLIRSYRHKRRGFNRKGIEMNKLSILPA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m108.pep	AILLGGCAAGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA					
g108	AILLGGCAAGG-NTFGSLDGGTGMGGSIVKMTVESQCRAELDRRSEWRLTALAMSAEKQA					
	70	80	90	100	110	
	130	140	150	160	170	
m108.pep	EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAALTAKTVSACFKHLYRX					
g108	EWENKICGCATEEAPNQLTGNDVMQMLNQSTRNQALAALT VKTVSACFKRLYRX					
	120	130	140	150	160	170

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 385>:

a108.seq

```

1  ATGTTGCCGG  GCTTCAACCG  GATATTCAAA  CGGTTTGTTT  CAACACTCGG
51  AACGGCGCAT  AAAACGCCGC  CCTTCGCGTT  ATCCCGAACG  GGGCGGCTAA
101 TCAGATTCTA  TCGCCATAAA  AGGCGGGGTT  TCAACCGAAA  AGGAATTGAG
151 ATGAATAAAA  CCTTGTCTAT  TTTGCCGGTG  GCAATCTTAC  TCGGCGGCTG
201 CGCCGCCGGG  GGCGGTAACA  CATTCCGGCAG  CTTAGACGGC  GGCACAGGTA
251 TGGCGGCGAG  CATCGTCAAA  ATGGCGGTAG  AAAGCCAATG  CCGTGCGGAA
301 TTGAACAAAC  GCAGCGAATG  GCGTTTGACC  GCGCTGGCGA  TGAGTGCCGA
351 AAAACAGGCG  GAATGGGAAA  ACAAGATTTG  CGCTTGCGTC  GCCCAAGAAG
401 CACCCAACCA  GCTGACCGGC  AACGATGTGA  TGCAGATGCT  GGATCCGTCC
451 ACGCGCAATC  AGGCACTTGC  CGCCCTGACC  GCCAAAACGG  TTTCCGCCTG
501 CTTCAAACAC  CTGTACCGCT  AA

```

This corresponds to the amino acid sequence <SEQ ID 386; ORF 108.a>:

a108.pep

```

1  MLPGFNRIK  RFVPTLGT  AH  K  T  P  P  F  A  L  S  R  T  G  R  L  I  R  F  Y  R  H  K  R  R  G  F  N  R  K  G  I  E  M  N  K  L  S  I  L  P  V
51  MNKTL  SILPV  AILLGGCAAG  GGNTFGSLD  G  T  G  M  G  G  S  I  V  K  M  A  V  E  S  Q  C  R  A  E
101 LNK  RSEWRLT  ALAMSAEKQA  EWENKICACV  AQEAPNQLTG  NDVMQMLDPS
151 TRNQALAALT  AKTVSACFKH  LYR*

```

m108/a108 96.5% identity in 173 aa overlap

	10	20	30	40	50	60
m108.pep	MLPGFNRIKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKLSILPV					
a108	MLPGFNRIKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKLSILPV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m108.pep	AILLGGCAAGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA					
a108	AILLGGCAAGGNTFGSLDGGTGMGGSIVKMAVESQCRAELNKRSEWRLTALAMSAEKQA					
	70	80	90	100	110	120
	130	140	150	160	170	
m108.pep	EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAALTAKTVSACFKHLYRX					
a108	EWENKICACVAQEAPNQLTGNDVMQMLDPSTRNQALAALTAKTVSACFKHLYRX					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 387>:

## g109.seq

```

1  ATGTATTATC GCCGGGTTGT GGGGCTATCC GATGGACTTG GCGATTTGGC
51 AGCCGGTATT GATCGTAGGC GTATGCTTAC CGCTTTTGGG AGCGGGCATG
101 GAAATGACGC GCAAAGGCAA AACCACCCAA TCCGCCGCCA TCGTGGTGTT
151 CTCTTCCGTC TGGTCAATCC GGTTCGCGC TGGGCGTTGA CGATGCTGTT
201 GGATAATTTG GGCTTAATCG GCTGCAAAGA ACGCAGCGCG CAATTAGGTT
251 TTGTCGGACG AGTATTGATA CCCGCAGTAG GTTCTTAAT CTTGTGTGTG
301 GCGATGGGTG CGGTCGGGAT GCTGCCCGGT ATCCCTCCGT TTTTGGAGCA
351 GTTCAAATCT TTGGGCTAG

```

This corresponds to the amino acid sequence <SEQ ID 388; ORF 109.ng>:

## g109.pep

```

1  MYYRRVVGLS DGLGDLAAGI DRRRMLTAFG SHHGNDQRQ NHPIRRHRGV
51  LFRVLNPFVG WALTMLLDNL GLIGCKERSA QLGFVGRVLI PAVGFLILCV
101  AMGAVGMLPG IPPFLEQFKS LG

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 389>:

## m109.seq

```

1  ATGTATTATC GCCGGGTTAT GGGGCTATCC GATGGACTTG GCGATTTGGC
51 AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCGCTTTTG
101 GAAGCGGGCA TGGAAATGAC GCGCAAAGGC AAAACCACCC AATCCGCCGC
151 CATCGTGGTG TTCTCTCCG CTTGTCAAT CCGGTTTTTCG GCTGGGCGTT
201 GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAA GAGCGCAGTG
251 CGCAATTAGG TTTCGCCGGA CGCGTGTGA TACCCGAGT AGGTTTCTTG
301 ATCTTGTGTG TGGCGATGGG TCGGTCGGG ATGCTGCCCG GTATCCGCC
351 GTTTTTGGAA CACTTCAAAT CTTTGGGCTA G

```

This corresponds to the amino acid sequence <SEQ ID 4; ORF 109>:

## m109.pep

```

1  MYYRRVMGLS DGLGDLAAGI ERSLGRRRIL TAFSGHGHND AQRONHPIRR
51  HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSAQLGFAG RVLIPAVGFL
101  ILCVAMGAVG MLPGIPPFLE HFKSLG*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 109 shows 92.9% identity over a 126 aa overlap with a predicted ORF (ORF 109.ng) from *N. gonorrhoeae*:

## m109/g109

```

              10      20      30      40      50      60
m109.pep    MYYRRVMGLSDGLGDLAAGIERSLGRRRILTAFGSGHGHND AQRONHPIRRHGVLFRLVN
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g109        MYYRRVVGSLDGLGDLAAGIDR---RRMLTAFSGHGHND AQRONHPIRRHGVLFRLVN
              10      20      30      40      50

              70      80      90     100     110     120
m109.pep    PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLI PAVGFLILCVAMGAVGMLPGIPPFLE
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g109        PVFGWALTMLLDNLGLIGCKERSAQLGFVGRVLI PAVGFLILCVAMGAVGMLPGIPPFLE
              60      70      80      90     100     110

m109.pep    HFKSLGX
              :||||
g109        QFKSLGX
              120

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 391>:

## a109.seq

```

1  ATGTATTATC GCCGGGTTGT GGGGCTATCC GATGGACTTG GCGATTTGGC
51 AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCGCTTTTG
101 GAAGCGGGCA TGGAAATGAC GCGCAAAGGC AAAACCACCC AATCCGCCGC
151 CACCGTGGTG TTCTCTCCG CTTGGTCAAT CCGGTTTTTCG GCTGGGCGTT

```

201 GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAA GAGCGCAGCG  
 251 CGCAATTAGG TTTCCACCGGA CGCGTATTGA TACCCGTAGT AGGTTTCTTG  
 301 ATCTTGTGTG TGGCGATGGG TGGGTGCGG ATGCTGCCCG GTATCCCGCC  
 351 GTTTTGGAG CACTTCAAAT CTTTGGGCTA G

This corresponds to the amino acid sequence <SEQ ID 392; ORF 109>:

a109.pep

1 MYYRRVVGLS DGLGDLAAGI ERSIGRRRIL TAFSGHGNDA AQRQNHPIRR  
 51 HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSQQLGFTG RVLIPVVGFL  
 101 ILCVAMGAVG MLPGIPPFLE HFKSLG\*

m109/a109 97.6% identity in 126 aa overlap

	10	20	30	40	50	60
m109.pep	MYYRRVMGLSDGLGDLAAGIERSLGRRRILTA	FGSGHGNDAQRQNHPIRRHRGVLFRLVN				
a109	MYYRRVVGLSDGLGDLAAGIERSLGRRRILTA	FGSGHGNDAQRQNHPIRRHRGVLFRLVN				
	10	20	30	40	50	60
	70	80	90	100	110	120
m109.pep	PVFGWALTMLLDNLGLIGCKERSAQLGFAGRV	LIPAVGFLILCVAMGAVGMLPGIPPFLE				
a109	PVFGWALTMLLDNLGLIGCKERSAQLGFTGRV	LIPVVGFLILCVAMGAVGMLPGIPPFLE				
	70	80	90	100	110	120
m109.pep	HFKSLGX					
a109	HFKSLGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 393>:

g111.seq

1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC  
 51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACG GCGCAAACCG  
 101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT  
 151 TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT  
 201 TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGtccACC TACCAGACCG  
 251 ATTCCGAAAT CAGCCGGTTt atacagacan atgctggaga gctcttcgcg  
 301 tntcatgcag nttctataac tgattccgcc gaagactgtc tgcctaatac  
 351 gcctatctca tcggcgctct ga

This corresponds to the amino acid sequence <SEQ ID 394; ORF 111.ng>:

g111.pep

1 MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLOGET MGTTYTVKYL  
 51 SNNRDKLPSP AKIQKRIDDA LKEVNRQMSY YQTDSEISRF IQTAGELFAH  
 101 ASITDSAEDC LPNTPISSAL \*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 395>:

m111.seq

1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGTCT TGATATTTGC  
 51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG  
 101 TTACCCTGCA AGGCGAAACG ATGGGCACGA CCTATAYCGT CAAATACCTT  
 151 TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AWAACGCAT  
 201 CGATGACGCG CTTAAAGAAK TCAACCGGyA GATGTCCACC TATCAGCCCG  
 251 ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC  
 301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC GCCTGAACCG  
 351 CCTGACACAC GGCGCGCTGG ACGTAACCGT CGGCCCTTGT GTCAACCTTT  
 401 GGGGATTTCGG CCCCACAA TCCGTTACCC GTGAACCGTC GCCGGAACAA  
 451 ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA  
 501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG  
 551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA



```

601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCGTGGCGC ATCGGTATCG
701 AGCAGCCCAA TATCGTCCAA GCGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAAACGTT CGCTTGCCAC TTCCGCGCAT TACCGTATTT TCCACGTCGA
801 TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCAGTG
901 ACGGCGGACG GCTTGTCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAAGCTCG TGTTCCTG ATTGTCAGGG
1001 ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTGA AAAAGTGC TC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 396; ORF 111>:

```

m111.pep
1  MPSETRLPNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTTYXVKYL
51  SNNRDKLPSP AEIXKRIDDA LKEXNRXMST YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
151 IKQAASYTGI DKIIILKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIVPL
251 NNRLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

ORF 111 shows 88.7% identity over a 97 aa overlap with a predicted ORF (ORF 111.ng) from *N. gonorrhoeae*:

m111.pep/g111.pep

```

              10      20      30      40      50      60
m111.pep      MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSP
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g111           MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
              10      20      30      40      50      60

              70      80      90      100     110     120
m111.pep      AEIXKRIDDALKEKNRXXMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTA EAVRLNRLTH
              |:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g111           AKIQKRIDDALKEVNROMSTYQTDSEISRFIQTXAGELFAXHAXSITDSAEDCLPNTPI
              70      80      90      100     110     120

              130     140     150     160     170     180
m111.pep      GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIIILKQKDYASLSKTHPK
g111           SALX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 397>:

a111.seq

```

1  ATGCCGCTCTG AAACACGCCT GCCGAAC TTT ATCCGCACCT TGATATTTGC
51  CCTGAGTTTT ATCTTCCTGA ACGCCTGTT C GGAACAAACC GCGCAAAACG
101 TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
201 CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
251 ACTCCGAAAT CAGCCGTTT C AACCAACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC ACCTGAACCG
351 CCTGACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCCG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAATCA TTTTGAAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCAAG GCCTATTTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGCGC ATCGGCATCG
701 AACAGCCCAA CATCGTCCAA GCGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAAACGTT CGCTTGCCAC TTCCGCGCAT TACCGTATTT TCCACGTCGA

```

801 TAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC  
 851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCATG  
 901 ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC  
 951 CTTAAAGCTG GCAGAGCGCG AAAAATCTGC TGTTCCTG ATTGTAGGG  
 1001 ATAAAGCGCG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC  
 1051 CGCTAA

This corresponds to the amino acid sequence <SEQ ID 398; ORF 111.a>:

a111.pep

1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLOGET MGTITYVKYL  
 51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMSY YQPDSEISRF NQHTAGKPLR  
 101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ  
 151 IKQAASYTGI DKIIKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE  
 201 LEKYGIQNYL VEIGGELHGK GKNARGEPRW IGIEQPNIVQ GGNTQIIVPL  
 251 NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPIHNLAS SISVVDASAM  
 301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL  
 351 R\*

m111/a111 97.7% identity in 351 aa overlap

m111.pep	10	20	30	40	50	60
	MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTITYXVKYLSNNRDKLPSP					
a111	:     :     :     :     :					
	10	20	30	40	50	60
	MPSETRLPNFIRTLIFALSFIFLNACSEQTAQTVTLQGETMGTITYVKYLSNNRDKLPSP					
m111.pep	70	80	90	100	110	120
	AEIXKRIDDALKEVNRXMSYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH					
a111	:     :     :     :     :					
	70	80	90	100	110	120
	AEIQKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVHLNRLTH					
m111.pep	130	140	150	160	170	180
	GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIIKQKGDYASLSKTHPK					
a111	:     :     :     :     :					
	130	140	150	160	170	180
	GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIIKQKGDYASLSKTHPK					
m111.pep	190	200	210	220	230	240
	AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPRWIGIEQPNIVQ					
a111	:     :     :     :     :					
	190	200	210	220	230	240
	AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPRWIGIEQPNIVQ					
m111.pep	250	260	270	280	290	300
	GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNNKRPIHNLASISVVDASAM					
a111	:     :     :     :     :					
	250	260	270	280	290	300
	GGNTQIIVPLNNRSLATSGDYRIFHVDKSGKRLSHIINPNNKRPIHNLASISVVDASAM					
m111.pep	310	320	330	340	350	
	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
a111	:     :     :     :					
	310	320	330	340	350	
	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 399>:

g111-1.seq

1 ATGCCGCTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTCG  
 51 CCTGGGTTTC ATCTTCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG  
 101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT  
 151 TCAAATAATC GGGACAAACT CCCCTCCCTT GCCAAAATAC AAAAGCGCAT  
 201 TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TACCAGACCG

251	ATTCCGAAAT	CAGCCGGTTC	AACCAACACA	CAGCCGGCAA	GCCCTCCGC
301	ATTTCAAGCG	ATTTCGACA	CGTACCGCC	GAAGCCGTCC	CGCTGAACCG
351	CCTGACTACG	GGTGCCTGG	ACGTAACCGT	GCGCCCTTTG	GCTCAACCTT
401	GGGGGTTCCG	CCCCGACAAA	TCCGTTACCC	GTGAACCGTC	GCCGGAACAA
451	ATCAAAACAGG	CGGCATCTTA	TACGGGACA	GACAAAATCA	TTTTCGAACA
501	AGGCAAAAGT	TACGCTTCCT	TGAGCAAAAC	TCACCCAAA	GCCTATTGG
551	ATTTATCTTC	GATTGCCAAA	GGCTTCGGCG	TTGATAAAGT	TGCGGGCGAA
601	CTGGA AAAAT	ACGGCATTTCA	AAATTATCTG	GTGAAAATCG	GcggcGAGTT
651	GCACGGCAAA	GGCAAAAATG	CGCACGGCGA	ACCGTGGCGC	ATCGGCTATAG
701	AGCAACCCAA	TATcatccaa	ggcggc caata	cgcAGattat	cgtcccgctg
751	aaCaaccggtt	cgcttgccac	ttccggcgAT	taccgtaTTT	tcacgctcgA
801	TAAAAACGGC	Aaacgccttt	cccacATCAT	CAATCCCAAC	AACAAAACGAC
851	CCATCAGcCA	CAAcctcgcc	tcCATCAgCg	TGGTCTCAGA	CAGTGC AATG
901	ACGGCGGACG	GGTTATCCAG	AGGATTATT	GTTT TAGCG	AACCCGAAGC
951	CTTAAGGCTG	GCAGAAACAAG	AAAAACTCGC	TGTTTTCTTA	ATTGTCGGGG
1001	ATAAGGACGG	CTACCGCACC	GCCATTGCTT	CCGAATTTCG	CAAGCTGCTC
1051	CGCTAA				

This corresponds to the amino acid sequence <SEQ ID 400; ORF 111-1.ng>:

g111-1.pap

1	MPSETRLPNL	IRALIFALGF	IFLNACSEQT	AQTVTLQGET	MGTTYTVKYL
51	SNNRDLKLSP	AKIQKRIDDA	LKEVNRQMS	YQTDSEISRF	NQHTAGKPELR
101	ISSDFAHVTA	EAVRLNRLTH	GALDVTGVL	VNLWGFGDPK	SVTREPSEKLR
151	IKQAASYTGI	DKIILLQOGK	YASLSKTHPK	AYLDDLSIAK	GFGVDKVAE
201	LEKYGQNYL	VEIGGELHGK	GKNAHGEPWR	IEIQEPNIIQ	GNNQITIVPL
251	NNRSLATSGD	YRIFHVDKNG	KRLSHIINPN	NKRPISHNLA	SISVVSDSAM
301	TADGLSTGLF	VLGETEARLR	AEQEKLAFLV	IVRDKDGYRT	AMSSEFAKLL
351	R*				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 401>:

m111-1.seq

1	ATGCCGCTCTG	AAACACGCCCT	GCCGAACTTT	ATCCGCGTCT	TGATATTTCG
51	CCTGGGTTC	ATCTTCTCTGA	ACGGCTGTTC	GGAACAAACC	GGCGAAACCG
101	TTACCTTGCA	AGGCGAAACG	ATGGGCACGA	CCTATACCGT	CAAAATACCTT
151	TCAAATTATC	GGGCAAAACT	CCCTGCACCT	GCGGAATATC	AAAACACGAT
201	CGATGACGCG	CTTAAAGAAG	TCAACCGGCA	GATGTCCACC	TATCAGCCCG
251	ACTCCGAAAT	CAGCCGGTTC	AACCAACACA	CAGCCGGGCA	GGCCCTCCGC
301	ATTTCAAGCG	ACTTCGCACA	CGTTACTGCG	GAAAGCGTCC	GCTCAACCG
351	CTTGACACAC	GGCGCGCTGG	AGCTAACCGT	CGGCCCTTG	GGCAACCTTT
401	GGGGATTTCG	CCCCGACAAA	TCCGTTACCC	GTGAACCGTC	GGCGGAACAA
451	ATCAAAACAG	CGCATCTTTA	TACGGGTCAC	GACAAAATCA	TTTTGAARAA
501	AGGCAAAAGT	TACGCTTCTT	TGAGCAAAAC	CACCCCAAG	GCGTTATTGG
551	ATTTATCTTC	GATTGCCAAA	GGCTTCGGCG	TTGATAAAGT	TGCGGGCGAA
601	CTGGAAAAAT	ACGGCATTTCA	AAATTATCTG	GTGGAATTCG	GGCGCGAGTT
651	GCACGGGCAA	GGCAAAAACG	CGCGGGCGCA	ACCGTGGCG	ATCGGTTATC
701	AGCAGCCCAA	TATCGTCCAA	GCGGGCAATA	CGCAGATTAT	GCTCCCGCTG
751	AACAACCGTT	CGTTTGCCAC	TTCGGCGGAT	TACCGTATTT	TCCACGTCGA
801	TAAAAACGCC	AAACGCCTCT	CCCATATCAT	CAACCCGAAC	AACAAACGAT
851	CCATACAGCA	CAACCTCGCC	TCCATCAGCG	TGGTCCGAGA	CAGTCGCGATG
901	ACGGCGGACG	GCTTGTCCAC	AGGATTATTC	GTATTGGGCG	AAACCGAAGC
951	CTTAAAGCTG	GCAGAGCGCG	AAAAACTCGC	TGTTTTCTTG	ATTGTCAGGG
1001	ATAAAGGCGG	CTACCGCACC	GCCATGTCTT	CCGAATTTGA	AAAAGTGCTC
1051	CGCTAA				

This corresponds to the amino acid sequence <SEQ ID 402; ORF 111-1>:

m111-1.pgp

1	MPSETRLPNI	IRVLIFALGF	<u>IFLNACSEQT</u>	AQTVTLQGET	MGTTYTVKYL
51	SNNRDKLPSP	AEIQKRIDDA	LKEVNMQST	YQPDSEISRF	NQHTAGSPLR
101	ISSDFAHVTA	EAVRLNRLTH	GALDVTVGPL	VNLWGFQDPK	SVTREPSEQR
151	IKQAASYTGI	DKIILKQKGD	YASLSKTHPK	AYLLDSSIAK	GFGVDKVAEG
201	LEKYGIQNYL	VEIGELHGK	GKNARGEPR	IGIEQPNIVQ	GGNTQIIIVL
251	NNRSLATSGD	YRIFHVDKNG	KRLSHIINPN	NKRPISHNLA	SISVVDASAM
301	TADGLSTGLF	VLGETEALKL	AEREKLAVFL	IVRDKGGYRT	AMSSEFEKLL
351	R*				

**m111-1/g111-1** 96.6% identity in 351 aa overlap

m111-1.pep MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPS  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
q111-1 MPSETRLPLRLIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPS

g111-1/p44550

sp|P44550|YOJL\_HAEIN HYPOTHETICAL LIPOPROTEIN HI0172 PRECURSOR >gi|1074292|pir||C64144  
hypothetical protein HI0172 - Haemophilus influenzae (strain Rd KW20) >gi|1573128 (U32702)  
lipoprotein, putative [Haemophilus influenzae Rd] Length = 346  
Score = 349 bits (885), Expect = 2e-95  
Identities = 177/328 (53%), Positives = 240/328 (72%), Gaps = 4/328 (1%)

Query:	23	LNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSPAEIXKRIDDALKEXNRXMSYQ	82
		L AC ++T + ++L G+TMGTTY VKYL + S + + I+ LK+ N MSTY+	
Sbjct:	17	LAACQKET-KVISLSGKTMGTTYHVKYLDGDSITATS-EKTHEEIEAILKDVNAKMSTYK	74
Query:	83	PDSEISRFNQHT-AGKPLRISDDFAHVTAEAVRLNRLTHGALDVTVGPLVNLWGFGPDKS	141
		DSE+SRFNQ+T P+ IS+DFA V AEA+RLN++T GALDVTVG+P+VNLWGFGP+K	
Sbjct:	75	KDSELSRFNQNTQVNTPIEISADFAKVLAEAIRLNKVTEGALDVTVGPPVNLWGFGPEKR	134
Query:	142	VTREPSPEQIKQAASYTGIDKIILKQGDYASLSKTHPKAYLDLSSIAGFGVDKVAGEL	201
		++P+PEQ+ + ++ GIDKI L K+ A+LSK P+ Y+DLSSIAGFGVD+VA +L	
Sbjct:	135	PEKQPTPEQLAERQAWVGIDKITLDTNKEKATLSKALPQVYVDLSSIAGFGVDQVAEKL	194
Query:	202	EKYGIQNYLVEIGGELHGKGNARGEPWRIGIEQPNIVQGGNTQIIVPLNNRSLATSGDY	261
		E+ QNY+VEIGGE+ KGKN G+PW+I IE+P + ++ LNN +A+SGDY	
Sbjct:	195	EQLNAQNYMVEIGGEIRAKGKNIEGKPWQIAIEKPTTTGERAVEAVIGLNNMGMASSGDY	254
Query:	262	RIFHVDKNGKRLSHIINPNNKRPISHNLASISVVADSAMTADGLSTGLFVLGETEALKLA	321
		RI+ ++NGKR +H I+P PI H+LASI+V+A ++MTADGLSTGLFVLGE +AL++A	
Sbjct:	255	RIY-FEENGKRFAHEIDPKTGYPQIHHLASITVLPAPTSMTADGLSTGLFVLGEDKALEVA	313
Query:	322	EREKLAVFLIVRDKGGYRTAMSSEFEKL	349
		E+ LAV+LI+R G+ T SS F+KL	
Sbjct:	314	EKNNLAVYLIIRTDNGFVTKSSSAFFKL	341

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 403>:

a111-1.seq

1	ATGCGGCTCTG	AAACACGCCT	GCCGAACTTT	ATCCGCACCT	TGATATTTCG
51	CCTGAGCTTTT	ATCTTCCTGA	ACGCTGTGTC	GGAAACAAAC	CGCGAAACCG
101	TTACAGCTGCA	AGGTGAAACG	ATGGGCACCA	CCTTATCCGT	CAAAATACCT
151	TCAAATATAC	GGGCAAAACT	CCCCTACCTT	CGCGAAATAC	AAAGCGCAT
201	CGATGCACGG	CTTAAAGAGT	TCAACCGGCA	GATGTCCACC	TATCAGCCCG
251	ACTCCGAAAT	CAGCCGGTTC	AACCAACACA	CAGCCGGCAA	GCCTCTCCGC
301	ATTTCAAGCG	ACTTCGCACA	CGTTACTGCC	GAAGCGCTCC	ACCTGAACCG

```

351 CCTGACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCGG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAAATCA TTTTGAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGCGC ATCGGCATCG
701 AACAGCCCAA CATCGTCCAA GCGGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAAACGTT CGCTTGCCAC TTCCGGCGAT TACCGTATT TCCACGTCGA
801 TAAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCATG
901 ACGGCGGACG GCTTGCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAGAGCTG GCAGAGCGCG AAAAATCTGC TGTTCCTG ATTGTCAGGG
1001 ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTGA AAAACTGCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 404; ORF 111-1.a>:

a111-1.pep

```

1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNQMST YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFDPK SVTREPSPEQ
151 IKQAASYTGI DKIIILKQGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPRW IGIEQPNIVQ GGNTQIIIVPL
251 NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

a111-1/m111-1 98.9% identity in 351 aa overlap

	10	20	30	40	50	60
a111-1.pep	MPSETRLPNFIRTLIFALSFIFLNACSEQTAQTVTLQGETMGTTYTVKYL	SNNRDKLPSP				
m111-1	MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYL	SNNRDKLPSP				
	70	80	90	100	110	120
a111-1.pep	AEIQKRIDDALKEVNQMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTA	EAVHLNRLTH				
m111-1	AEIQKRIDDALKEVNQMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTA	EAVRLNRLTH				
	130	140	150	160	170	180
a111-1.pep	GALDVTVGPLVNLWGFDPKSVTREPSPEQIKQAASYTGIDKIIILKQGD	YASLSKTHPK				
m111-1	GALDVTVGPLVNLWGFDPKSVTREPSPEQIKQAASYTGIDKIIILKQGD	YASLSKTHPK				
	190	200	210	220	230	240
a111-1.pep	AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPRW	IGIEQPNIVQ				
m111-1	AYLDLSSIAKGFGVDKVAGELEKYGIQNYLVEIGGELHGKGKNARGEPRW	IGIEQPNIVQ				
	250	260	270	280	290	300
a111-1.pep	GGNTQIIIVPLNNRSLATSGDYRIFHVDKSGKRLSHIINPNKRPISHNLA	SISVVADSAM				
m111-1	GGNTQIIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNKRPISHNLA	SISVVADSAM				
	310	320	330	340	350	
a111-1.pep	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
m111-1	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 405>:

g114.seq

```

1 ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCAGCAGG AATGCAGCAA
51 GACTTTTTTA TGTCCGCCGG GCGGGACGAG TATGGGGCGG TCAATGTCGG

```

```

101 TAACGGTAGG TTTGTTTTGT GTTTCATTA ACTTAACAAT ATCTGTCGAA
151 TACGGTCAAA GCGGCTATTT TACCAGAGCC GCCGAATGTA AAACAGGGTG
201 TCAGGGCATC AGCCCGAGCT GCCTGAACGA ACGGACGGTT TCGGAGGTAA
251 CGATAAAATG GTCGAGCAGC GAAACATCAA CCAGCGACAT GGCCTGTGCC
301 AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAAGGTTTCTG GCGAGCCGCC
351 CGGATGGTTG TCGCGGATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
401 GTTTGACGAT TTCGCGGATG TAA

```

This corresponds to the amino acid sequence <SEQ ID 406; ORF 114.ng>:

```

g114.pep
  1 MASITSPLHG AQQECSTFL CPPGGTSMGR SMSVTVGLFC VSINLTISVE
 51 YGQSGYFTRA AECKTGCGI SPSCLNERTV CEVTIKWSSS ETSTSDMACA
101 SRLVNMSSC EGSGEPPGWL CAIIRLSAYS SNASLTISRM *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 407>:

```

m114.seq
  1 ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCACAGAG AATGCAGCAA
 51 GACTTTTTTA TGTCCACCGG GCGGGACGAG TATAGGGCGG TCAATGTCGG
101 TAACGGTAGG TTTGTTTTGT GTTTCATTA ACTTAACAAT ATCTGTTGAA
151 TACGGTTGAA GCGGCTATTT TATCAGAGCC GCCGCATGTA AAACAGAGTG
201 TCAGGGCATC AACCCGAGCT GTCTGAACGA ACAGACGCTT TCGGAGGTAA
251 CGATAAAATG GTCGAGCAGC GACACATCGA CCAGCGACAT TGCCTGTGCC
301 AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAAGGTTCTG GCGAGCCGCC
351 CGGATGGTTG TCGGCAATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
401 GTTTGACGAT TTCGCGGATG TAA

```

This corresponds to the amino acid sequence <SEQ ID 408; ORF 114>:

```

m114.pep
  1 MASITSPLHG AHRECSKFL CPPGGTSIGR SMSVTVGLFC VSINLTISVE
 51 YGXSGYFIRA AACKTECQGI NPSCLEQTL CXVTIKWSSS DTSTSDIACA
101 SRLVNMSSC EXSGEPPGWL CAIIRLSAYS SNASLTISRM *

```

m114/g114 90.0% identity over a 140 aa overlap

	10	20	30	40	50	60
m114.pep	MASITSPLHGAHRECSKFLCPPGGTSIGRSMSTVGLFCVSINLTISVEYXSGYFIRA					
	10	20	30	40	50	60
g114	MASITSPLHGAQQECSTFLCPPGGTSMGRSMSTVGLFCVSINLTISVEYQSGYFTRA					
	70	80	90	100	110	120
m114.pep	AACKTECQGINPSCLEQTLCXVTIKWSSSDTSTSDIACASRLVNMSSCEXSGEPPGWL					
	70	80	90	100	110	120
g114	AECKTGCGISPSCLNERTVCEVTIKWSSSETSTSDMACASRLVNMSSCEGSGEPPGWL					
	130	140				
m114.pep	CAIIRLSAYSSNASLTISRMX					
	130	140				
g114	CAIIRLSAYSSNASLTISRMX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 409>:

```

a114.seq
  1 ATGCCGGAGG CAAGCATCGC CTCCATCACT TCGCCGCTGC ACGGGGCGCA
 51 ACAGGAATGC AGCAAGACTT TTTTATGTCC GCCGGGCGGG ACGAGTATGG
101 GCGGGTCAAT GTCGGTAACG GTAGGTTTGT TTGTGTTTC CATTAACCTA
151 ACGATATCTG TCGAATACGG TTGAAGCGGC TATTTTATCA GAGCCGCCGC
201 ATGTAAACA GGGTGTGAGG GCATCAGCCC GAGCTGCCTG AACGAACGGA
251 CGGTTTGC GCATTACGATA AAATGGTCTG GCAGCGACAC ATCGACCAGC
301 GACATTGCCT GTGCCAGCCG CCTGTGAAC ATGATGTCTT CCTGCGAAGG
351 TTCGGGCGAG CCGCCCGGAT GGTGTGCGC GATAATCAGG CTGTCGGCAT
401 ATTCGTCAA TGCCAGTTTG ACAATTTTAC GGATGTAA

```

This corresponds to the amino acid sequence <SEQ ID 410; ORF 114.a>:

a114.pep

```

1  MPEASIASIT SPLHGAQQEC SKTFLCPPGG TSMGRMSVTVGLFCVSINL
51  TISVEYG*SG YFIRAAACKT GCQGISPSCL NERTVCAVTI KWSSSDTSTS
101 DIACASRLVN MMSSCEGSGE PPGWLCIIR LSAYSSNASL TISM*
```

m114/a114 92.9% identity in 140 aa overlap

```

              10      20      30      40      50
m114.pep      MASITSPLHGAHRECSKTFLCPPGGTSMGRMSVTVGLFCVSINLTISVEYGXSG
              :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a114           MPEASIASITSPLHGAQQECSTFLCPPGGTSMGRMSVTVGLFCVSINLTISVEYGXSG
              10      20      30      40      50      60

              60      70      80      90      100     110
m114.pep      YFIRAAACKTECQGINPSCLEQTLXVTIKWSSSDTSTSDIACASRLVNMSSCEXSSE
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a114           YFIRAAACKTGCQGISPSCLNERTVCAVTIKWSSSDTSTSDIACASRLVNMSSCEGSSE
              70      80      90      100     110     120

              120     130     140
m114.pep      PPGWLCIIRLSAYSSNASLTISM*
              |||||:|||||:|||||:|||||
a114           PPGWLCIIRLSAYSSNASLTISM*
              130     140
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 411>:

g117.seq

```

1  atgggtcgacg aactcgacCT GCTGCCCCGAT GCCGTCGCCG CCACCCTGCT
51  TGCCGACATC GGACGCTACG TCCCCGATTG GAACCTATTG GTTTCGAGC
101 GCTGCAACAG CACCGTCGCC GAGCTGGTCA AAGGTGtgga CGAAGTGCAG
151 AAACCTTACCC ACTTCGCCCG GGTGGACAGC CTCGCCACGC CGGAAGAACG
201 CGCACAGCAA GCGGAAACCA TCGGAAAAAT GCTGCTGGCg atggttaccg
251 Acatccgcgt cgtaTTAATC AAACCTGGCGA TGCGTAcgcg caccCTGcta
301 tTTTtaaGCA ACGCCCAAGCA CAGCCCTGAA AAACgcgccG TCgcaaaAga
351 aacccTCGAC ATCTTCGCCC CGCTCGCCA CCGCTTGGGC GTGTGGCAGC
401 TCAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
451 TACCGCGAAA TCGCCCTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
501 ATACATCGAA AACTTCCTCG ATATCCTGCG TACGGAACCTC AAAAAATACA
551 ATATCCACTT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTAC
601 AAAAAAATGG TGAAGAAAAA ACTCAGCTTC GACGgccTGT TCGACATCCG
651 CGCCGTGCGG ATTCTGGTCG ATACCGTCCC CGaGTGTtAC ACCACGCTGG
701 gcaTCGTCCA CAGCCTCTGG CAGCCCATTC CCGGCGagtt CGAcgactAC
751 ATCGCCAACC CCAAAGgcaA CGgttATAAA AGtTTGCACA CCGTCATCGT
801 cggcccGGAa gacaaaggtg tggaaGtgCA AATCCGCACC TTCGAtatGC
851 accAATTcAa CgaatTcggT gtcgcccGCC ACTGGCGtta caaagaaggc
901 ggcaaaggcg attccGCCTA cgaacaaAAA ATcgccTggt TGCgccaACT
951 CTTGGACTGG CGCGAAAATA TGGCGGAAAG CGGCAAGGAA GACCTCGCCG
1001 CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTtATGT TTTGACCCCG
1051 CACGGCAAAG TCCTCTCTCT GCCAACGGGC GCAACCCCA TCGACTTCGC
1101 CTACGCCCTG CACAGCAGCA TcggCGACCG CTGCCGGGGC GCGAAAGTCG
1151 AaggGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACGG ACAGCGCGTC
1201 GAAATcatta cCGCcaaAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
1251 AGGctgGGtc aaATCCGGCA AGGCCATCGG caaAATCCCG GCCTAcatCC
1301 GCCAGcaaAa cgCgaCACC GTGCGCGAAG AAGGCCGTGT CCAACTCGAC
1351 AAGCAGCTTG CCAAACCTAC GCCCAAACCC AACCTGCAAG AGCTTgcca
1401 aaATCTCGGC taaaaAAAGC cagaagacct ctacacCGCc gtcggacaag
1451 gcgaaatttc caaccgcgcc atCaaaaaag cctgcggcac GCTgaacgaa
1501 ccgccccCGG TGCCCGTCAG CGCAACCACC ATCGTCAAAC AGTCCAAAT
```

```

1551 CAAAAAAGGT GGCAAAACCG GCGTGCTCAT CGACGGCGAA GACGGCTTGA
1601 TGACCACGCT TGCCAAATGC TGCAAACCCG CGCCGCCCGA CGATATTGCC
1651 GGCTTCGTTA CCCGCGAGCG CGGCATTTCG GTCCACCGCA AAACCTGCCC
1701 CTCTTCCGA CACCTTGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1751 GTTGGGCGGC GTTGCAGGAA GGGCAAGTGT TCGCCGTCGA TATCGAAATC
1801 CGGCCCAAG ACCGCTCCGG GCTTTGCGC GACGTATCCG ACGCGCTCGC
1851 CCGCCACAAA CTCAACGTTA CCGCCGTGCA AACCAGTCC CGGACTTGG
1901 AAGCCAGCAT GAGGTTACG CTCGAAGTCA AACAAGtCAA CGacCTCCCG
1951 CCGTCCTCG CCGCCTCGG CGATGTCAA GCGTATTGA GCGTTACCCG
2001 GCTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 412; ORF 117.ng>:

g117.pep

```

1 MVDEL DLLPD AVAATLLADI GRYVPDWNLL VSERCNSTVA ELVKGVDEVO
51 KLTHFARVDS LATPEERAQQ AETMRKMLLA MVTDIRVLI KLAMRTRTLL
101 FLSNAPDSPE KRAVAKETLD IFAPLANRLG VWQLKWQLED LGFRHQEPEK
151 YREIALLLDE KRTERLEYIE NFLDILRTEL KKYNIHFEVA GRPKHIYSIY
201 KKMVKKKLSF DGLFDIRAVR ILVDTVPECY TTLGIVHSLW QPIPGFDDY
251 IANPKGNGYK SLHTVIVGPE DKGVVQIRT FDMHQFNEFG VAAHWRYKEG
301 KGDSAYEQK IAWLRQLLDW RENMAESGKE DLAAAFKTEL FNDTIYVLT
351 HGKVLSLPTG ATPIDFAYAL HSSIGDRCRG AKVEGQIVPL STPLENGQRV
401 EIITAKEGHP SVNWLYEGWV KSGKAIGKIR AYIROQNADT VREEGRVQLD
451 KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
501 PPPVPVSATT IVKQSKI KKG GKTGVLIDGE DGLMTTLAKC CKPAPPDDIA
551 GFVTRERGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI
601 RAQDRSGLLR DVSDALARHK LNVTAQTQS RDLEASMRFT LEVKQVNDLP
651 RVLAGLGDVK GVLSVTRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 413>:

m117.seq (partial)

```

1 ..GTGAAACTCA AGAAATACAA TGTCCATTTC GAAGTCGCCG GCCGCCCGAA
51 ACACATCTAC TCCATTTACA AAAAAATGGT GAAGAAAAAA CTCAGCTTCG
101 ACGGCCTCTT TGACATCCGC GCCGTGCGAA TTCTGGTTGA TACCGTCCCC
151 GAGTGTACAC CCACGCTGGG TATCGTCCAC AGCCTCTGGC AGCCCATTC
201 CGGCGAGTTC GACGACTACA TCGCCAATCC CAAAGGCAAC GGCTATAAAA
251 GTTTGCACAC CGTCATCGTC GGCCCGGAAG ACAAAGGCGT GGAAGTACAA
301 ATCCGCACCT TCGATATGCA CCAATTCAAC GAATTCCGGT TCGCCGCCCA
351 CTGgCGTTAC AAAGAGGGCG GCAAGGGCGA TTCCGCCTAC GAACAGAAAA
401 TCGCCTGGTT GCGCCAATC TTGGACTGGC GCGAAAACAT GGCGGAAAGC
451 GGCAAGGAAG ACCTCGCCGC CGCCTTCAA ACCGAGCTTT TCAACGACAC
501 GATTTATGTT TTGACCCCGC ACGGCAAAGT CCTCTCCCTG CCCACGGGCG
551 CGACCCCAT CGACTTCGCC TACGCCCTGC ACAGCAGCAT CGGCGACCGT
601 TGCCGCGGTG CGAAAGTCGA AGGGCAGATT GTGCCGCTGT CCACCCCGCT
651 CGAAAACGGA CAGCGCGTCG AAATCATTAC CGCCAAAGAA GGGCATCCTT
701 CCGTCAACTG GCTTTACGAA GGCTGGGTCA AATCCAACAA GGCAATCGGC
751 AAAATCCGCG CCTACATCCG CCAGCAAAAC GCCGACACCG TGCGCGAAGA
801 AGGCCGCGTC CAACTCGACA AACAGCTTGC CAAACTCAGC CCCAAACCCA
851 ACCTGCAAGA GCTTGCCGAA AATCTCGGCT ACAAAGGCC AGAAGACCTC
901 TACACCGCCG TCGGACAAGG CGAAATTTC AACC CGCCA TCCAAAAAGC
951 CTGCGGCACg CTGAACGAAC CGCCGCCCGT ACCCGTCAGC GAAACCCACA
1001 TCGTCAAACA GTCCAAAATC AAAAAAGGCG GCAAAAACGG CGTGCTCATC
1051 GACGGCGAAG ACGGTCTGAT GACCACGCTT GCCAAATGCT GCAAACCCGC
1101 GCCGCCGAC GATATTATCG GCTTCGTTAC CGCGAGCGC GgCATTTCAG
1151 TGCACCGCAA AwyyTkCyCG TCTTTCCAAC ACCTCGCCGA ACACGCGCCC
1201 GAWAAAGTGC TGGACGCAAG CTGGGCGGCA TTGACGGAAG GACAAGTATT
1251 CGCCGTCGAT ATCGAAATCC GCGCCCAAGA CCGCTCCGGG CTTTTCGCGC
1301 ACGTATCCGA CGCGCTCGCC CGCCACAAAC TCAACGTTAC CGCCGTGCAA
1351 ACCCAGTCCC GCGACTTGA AGCCAGCATG AGGTTACGC TCGAAGTCAA
1401 ACAAGTCAAC GACCTCCCGC GCGTCTCGC CAGCCTCGGC GACGTCAAAG
1451 GCGTATTGAG CGTACCCGG CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 414; ORF 117>:

m117.pep (partial)



Homology with a predicted ORF from *N. gonorrhoeae*

ORF 117 shows 97.6% identity over a 490 aa overlap with a predicted ORF (ORF 117.ng) from *N. gonorrhoeae*:

m117/q117

m117.pep VKLKKYNVHFEVAGRPKHIYSIYKKMVKKKL  
 :|||:|||||||  
 g117 EKYREIALLLDEKRTRELEYIENFLDILRTCLKKYNIHFEVAGRPKHIYSIYKKMVKKKL  
 150 160 170 180 190 200

40 50 60 70 80 90  
 m117.pep SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG  
 |||||  
 g117 SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG  
 210 220 230 240 250 260

100 110 120 130 140 150  
 m117.pep PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG  
 g117 PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG  
 270 280 290 300 310 320

```

      160      170      180      190      200      210
m117.pep  KEDLAAAFKTELNDTIYVLTTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV
          |||||
g117      KEDLAAAFKTELNDTIYVLTTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV
          |||||
      330      340      350      360      370      380

```

m117.pep  
g117

PLSTPLENGQ RVEIIITAKEGHPSVNWL YEGVWKSNKAIGKIRAYIRQQNADTVREEGRVQ  
PLSTPLENGQ RVEIIITAKEGHPSVNWL YEGVWKSGKAIGKIRAYIRQQNADTVREEGRVQ

220 230 240 250 260 270  
390 400 410 420 430 440

280            290            300            310            320            330  
 m117.pep    LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPPVPVSE  
              |||||  
 g117        LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPPVPVSA  
              |||||  
 450            460            470            480            490            500

```

      340      350      360      370      380      390
m117.pep  TTIVQSKIKKGGKNGVLIDGEDGLMTTLAKCKPAPPDDIIGFVTRERGISVHRKXXXS
           |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g117      TTIVQSKIKKGGKTGVLIDGEDGLMTTLAKCKPAPPDDIAGFVTRERGISVHRKTCPS
           510      520      530      540      550      560

```

400 410 420 430 440 450  
m117.pep FQHLAEHAPXKVLDSWAALQEGQVFAVDIEIRAODRSGLLRDVSDALARHKLNVTAVOT

[illegible]

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 415>:

a117.seq

1	ATGTTTCATG	AACTCGACCT	GCTCCCGCAT	GCGTCGCGG	CCACCTTGCT
51	TGCCGACATC	GGACGCTACG	TCCCGCATG	GAACCTATTG	GTTTCCGAAC
101	GCTGCAACAG	TACCGTCGCC	GAGCTGGTCA	AAGGTGTGGA	CGAAGTGCAG
151	AAACTCACCC	ACTTCGCCCC	GGTGACAGC	CTCGCCACGC	CGGAAGAACG
201	CGCCACGACG	CGAGAAACTA	TGCGGAAAAA	GCTGCTGGCG	ATGGTTACCG
251	ACATCCGCGT	CGTGTTAATC	AAACTGGCGA	TGCTGACGCG	CACCTTCGAA
301	TTTTTAAGCA	ACGCCCCCGA	CAGCCCCGAA	AAACGCGCCG	TCGCGAAAGA
351	AACCTCGGAC	ATCTTCGCCC	CGCTCGCCAA	CCGTTTGGGC	GTGTGGCAGC
401	TCAAATGGCA	GCTCGAAGAT	TGGGGCTTCC	GCCATCAAGA	ACCCGAAAAA
451	TACCGCGAAA	TCCGCTTGCT	TTTGACGAGA	AAACGCAACG	AACGCTTCGA
501	ATACATCGAA	AACTTCCTTA	ATATCCTGCG	TACGGAACCT	AAAAAATACA
551	ATATCCACTT	TGAAGTCGCC	GGCCGTCCGA	AACACATCTA	CTCCATTTAC
601	AAAAAATGG	TGAAGAAAAA	ACTCAGCTTC	GACGGGTTGT	TCGACATCCG
651	CGCGTGC GG	ATTCTGGTTG	ATACCGTCCC	CGAGTGTATT	ACCACACTGG
701	GCATTGTCCA	CAGCCTCTGG	CAGCCCATTC	CGCGCGAGTT	CGACGACTGC
751	ATCGCCAACC	CGAAAGGCAA	CGGCTATAAA	AGTTTGCACA	CGGTCATCGT
801	CGGCCCGGAA	GACAAAGGCG	TGGAAGTGCA	AATCCGCACC	TTCGATATGC
851	ACCAAATGCA	CGAATTCCGT	GTCGCCGCGC	ACTGCGCTTA	CAAAGAGGGC
901	GGCAAAGGCG	ATTCCGCTTA	CGAAACAAAA	ATCGCTTGTT	TACGCCAACT
951	TTTGGA CTGG	CGCGAAAAACA	TGGCGGAAAG	CGGCAAGGAA	GACCTCGCGG
1001	CCGCCTTCAA	AACCGAGCTT	TTCAACGACA	CGATTTATGT	TTTGACCCCG
1051	CACGGCAAAG	TCCTCTCCCT	GCCCAACGGC	GCGACCCCCA	TCGACTTCGG
1101	CTACGCCCTG	CACAGCAGCA	TCCGCGACCG	TTGCCGCGGT	GCGAAAGTGC
1151	AAGGGCAGAT	TGTGCCGCTG	TCCACCCCGC	TCGAAAACGG	ACAGCGTGTG
1201	GAAATCATT A	CCGCCAAAGA	AGGGCATCCT	TCCGTCAACT	GGCTTTACGA
1251	AGGCTGGGTC	AAATCCAACA	AGGCAATCGG	CAAAATCCGC	GCCTACATCC
1301	GCCAGCAAAA	CGCCGACACC	GTGCGCGAAG	AAGGCCGCGT	CCAACTCGAC
1351	AAACAGCTTG	CGAAACTCAC	GCCCAAAACC	AACCTGCAAG	AGCTTGCCGA
1401	AAATCTCGGC	TACAAAAAGC	CAGAAGACCT	CTACACCGCC	GTCGGACAAG
1451	GCGAAATTTT	CAACCGCGCC	ATCCAAAAAG	CCTGCGGCAC	GCTGAACGAA
1501	CCGCGCGGCG	TACCCGTGAC	CGAAATCCACC	ATCGTCAAAC	AGTCCAAAAT
1551	CAAAAAGCCG	GGCAAAAACG	GCGTGCTCAT	CGACGGCGAA	GACGGTCTGA
1601	TGACCACGCT	TGCCAAATGC	TGCAAAACCG	CGCCGCCCGA	CGACATTGTC
1651	GGCTTCGTTA	CCCGCGATCG	CGGCATTTTC	GTACACCGCA	AAACTGCGCC
1701	CTCTTTCCGA	CACCTCGCCG	AACACGCGCC	CGAAAAAGTA	CTGGACGCAA
1751	GTTGGCGCGC	GTTGCAGGAA	GGACAAGTGT	TCGCGCTCGA	TATCGAAATC
1801	CGCGCCCAAG	ACCGCTCCGG	GCTTTTGGCG	GACGTATCCG	ACGCGCTCGC
1851	CCGCCACAAA	CTCAACGTTA	CCGCCGTGCA	AACCATGTC	CGCGACTTGG
1901	AAGCGACGAT	GAGGTTACAG	CTCGAAGTCA	AACAAGTTAC	CGACCTCCCA
1951	CGCGTCTCTG	CCAGCCTCGG	CGACGTCAAA	GGCGTATTGA	GCGTTACCCG
2001	GCTTTAA				

This corresponds to the amino acid sequence <SEQ ID 416; ORF 117.a>:

a117.pap

1	MVHELDLLPD	AVAATLLADI	GRYVPDWNLL	VUSERCNSTVA	ELVKGVDEVO
51	KLTHFARVDS	LATPEERAQQ	AETMRKMLLA	MVTDIRVVLI	KLAMRTRTLQ
101	FLSNAPDSPE	KRAVAKETLD	IFAPLANRLG	VWQLKWQLED	LGFRHQEPEK
151	YREIALLLDE	KRTERLEYIE	NFLNILRTEL	KKYNIHFEVA	GRPKHIYSIY
201	KKMWKKLLSF	DGLFDIRAVR	ILVDTVPECY	TTLGIVHSLW	QPIPGFEDDY
251	IANKPKNGYK	SLHTVIVGPE	DKGVEQIIRT	FLMHQFNEFG	VAAHWRYKEG
301	KGKDSAYEQK	IAWLRQLLDW	RENMAESGKE	DLAAMFKTEL	FNDTIYVLTP
351	HGKVLSLPTG	ATPIDFAYAL	HSSIGDRCRG	AKVEGQIVPL	STPLENGQRV

[illegible]

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 417>:

g117-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CCCTGCAAGA
51  ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCGG GACAACGATA
101  AAAACCTCAT CGGTACCGCA TGGTCTGCTG CGCAGGAACA TTATCCTGCC
151  GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
201  GGCAGCAATG GTCGACGAAC TCGACCTGCT GCCCGATGCC GTCGCCGCCA
251  CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGATTGGAA CCTATTGGTT
301  TCCGAGCGCT GCAACAGCAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351  AGTGCAGAAA CTTACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401  AAGAACGCGC ACAGCAAGCG GAAACCATGC GGAAAAATGCT GCTGGCGATG
451  GTTACCGACA TCCGCGTCGT ATTAATCAAA CTGGCGATGC GTACGCGCAC
501  CCTGCAATTT TTAAGCAACG CCCCCGACAG CCCTGAAAAA CGCGCCGTCG
551  CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG CTTGGGCGTG
601  TGGCAGCTCA AATGGCAGCT CGAAGATTTC GGCTTCGGCC ATCAAGAACC
651  CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
701  GCCTCGAATA CATCGAAAAA TCCTCGATA TCCTGCGTAC GGAACCTCAA
751  AAATACAATA TCCACTTTGA AGTCGCGCGC CGTCCGAAAC ACATCTACTC
801  CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTGTTCG
851  ACATCCGCGC CGTCCGGATT CTGGTCGATA CCGTCCCCGA GTGTACACC
901  ACGCTGGGCA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGagttCGA
951  cgactACATC GCCAACCCCA AAGgcaACGg ttATAAAAGT TTGCACACCG
1001 TCATCGTcgg cccGGAagaa aaaggtgtg gagtgcAAAT CCGCACCTTC
1051 GATATGCacc AATTCAaCga ATTcGGTGTc GCCGCCCACT GGCGTTACAA
1101 AGAAGCGCGC AAAGCGGATT CCGCCTACGA ACAAAAAATC GCCTGGTTGC
1151 GCCAACTCTT GGACTGGCGC GAAAATATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCGCCG CTTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGCAC GGCAAGTCC TCTCTCTGCC AACGGGCGCA ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGCTG CCGGGGCGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCGCGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAA TCCGGCAAGG CCATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGTGTCCA
1551 ACTCGACAAG CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTGC CCGTCAGCGC AACCACCATC GTCAAACAGT
1751 CCAAATCAA AAAAGGTGGC AAAACCGGCG TGCTCATCGA CGCGAAGAC
1801 GGCTTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851 TATTGCCGCG TTCGTTACCC GCGAGCGCGG CATTTCCGTC CACCGCAAAA
1901 CTTGCCCTTC TTTCCGACAC CTTGCCGAAC ACGCGCCCGA AAAAGTACTG
1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGG CAAGTGTTTC CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAACTC AACGTTACCG CCGTGCAAA CAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
2151 CCTCCGCGC GTCTCGCCG GCCTCGCGA TGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTA

```

This corresponds to the amino acid sequence <SEQ ID 418; ORF 117-1.ng>:

g117-1.pep

```

1  MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WSLAQEHYPA
51  DAATPYGEPL PDHFLGAAQM VDELDLDPDA VAATLLADIG RYVPDWNLLV
101  SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151  VDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201  WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLDILRTELK
251  KYNHFVAVAG RPKHIYSIYK KMKKKLSFD GLFDIRAVRI LVDTVPECYT
301  TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPEE KGVEVQIRTF
351  DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401  LAAAFKTELF NDTIYVLTPL GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451  KVEGQIVPLS TPLENGORVE IITAKEGHPV VNWLYEGWVK SGKAIGKIRA
501  YIRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551  GQGEISNRAI QKACGTLNEP PPVPVSATTI VKQSKIKKGG KTGVLIDGED
601  GLMTTLAKCC KPAPPDDIAG FVTRERGIVS HRKTCPSFRH LAEHAPEKVL
651  DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701  DLEASMRFTL EVKQVNDLPR VLAGLGDVKG VLSVTRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 419>:

m117-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA

```

```

51  ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
101 AAAACCTCAT CGGTACCGCA TGGTTGCTGG CGCAGGAACA TTACCCCGCC
151 GATGCCGCGC CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
201 GCGCGAAATG GTTCATGAAC TCGACCTGCT CCCCAGTGCC GTCCGCCCA
251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
301 TCCGAACGCT GCAACAGTAC CGTCGCGGAG CTGGTCAAAG GTGTGGACGA
351 AGTGCGAGAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401 AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAAATGCT GCTGGCGATG
451 GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501 CCTGCAATTT TTAAGCAACG CCCCCGACAG CCCCAGAAAA CGCGCCGTCG
551 CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG TTTGGGCGTG
601 TGGCAGCTCA AATGGCAGCT CGAAGATTTG GGCTTCCGCC ATCAAAGACC
651 CGAAAAATAC CGCGAAATCG CGCTGCTTTT GGACGAAAAA CGCACCGAAC
701 GCCTCGAATA CATCGAAAAA TTCCTCAACA TCCTGCGCGG TGAACCTAAG
751 AAATACAATG TCCATTTCGA AGTCGCGGCG CGCCCCGAAAC ACATCTACTC
801 CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTCTTTG
851 ACATCCGCGC CGTGCGAATT CTGGTTGATA CCGTCCCGCA GTGTTACACC
901 ACGCTGGGTA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
951 CGACTACATC GCCAATCCCA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001 TCATCGTCGG CCGGAAGAC AAAGGCGTGG AAGTACAAAT CCGCACCTTC
1051 GATATGCACC AATTCAACGA ATTGGGTGTC GCCGCCCACT GGCGTTACAA
1101 AGAGGCGCGC AAGGCGGATT CCGCCTACGA ACAGAAAAAT GCCTGGTTGC
1151 GCCAACTCTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCGCGCG CTTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGCAC GGCAAAGTCC TCTCCCTGCC CACGGGCGCG ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCGCTCG AAAACGGACA
1401 CGCGCTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1551 ACTCGACAAA GAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751 CCAAAATCAA AAAAGGCGGC AAAAAGCGCG TGCTCATCGA CGCGGAAGAC
1801 GGTCGTGATG CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851 TATTATCGGC TTCGTTACCC GCGAGCGCGG CATTTCAGTG CACCGCAAAA
1901 CCTGCGCCGC TTTTCAACAC CTCGCGGAAC ACGCGCCCGA AAAAGTGCTG
1951 GACGCAAGCT GGGCGGCATT GCAGGAAGGA CAAGTATTCG CCGTCGATAT
2001 CGAAATCCCG GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
2151 CCTCCGCGCG GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTAA

```

This corresponds to the amino acid sequence <SEQ ID 420; ORF 117-1>:

#### m117-1.pep

```

1  MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WLLAQEHYPA
51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQKPEKY REIALLLDEK RTERLEYIEN FLNLRGELK
251 KYNVHFEVAG RPKHIYSIYK KMKKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401 LAAAFKTELF NDTIYVLTTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQORVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501 YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 GQGEISNRAI QKACGTLNKP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601 GLMTTLAKCK KPAPPDDIIG FVTRERGISV HRKTCPSFQH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRFTL EVKQVNDLPR VLASLGDKVG VLSVTRL*

```

m117-1/g117-1 98.2% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSATLQELREWFDYSYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL					
g117-1	MTAISPIQDTQSATLQELREWFDYSYCAALPDNDKNLIGTAWSLAQEHYPADAATPYGEPL					
	10	20	30	40	50	60
	70	80	90	100	110	120

m117-1.pep PDHFLGAAQMVHEDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVEVQK  
g117-1 PDHFLGAAQMVDELDDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVEVQK  
70 80 90 100 110 120

m117-1.pep 130 140 150 160 170 180  
LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK  
g117-1 LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK  
130 140 150 160 170 180

m117-1.pep 190 200 210 220 230 240  
RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIEN  
g117-1 RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN  
190 200 210 220 230 240

m117-1.pep 250 260 270 280 290 300  
FLNILRGELKKYNVHFEVAGRPKHIYSIYKKMVKKLSFDGLFDIRAVRILVDTVPECYT  
g117-1 FLDILRTELKKYNIHFEVAGRPKHIYSIYKKMVKKLSFDGLFDIRAVRILVDTVPECYT  
250 260 270 280 290 300

m117-1.pep 310 320 330 340 350 360  
TLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTQFDMHQFNEFGV  
g117-1 TLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVGPEEKGVEVQIRTQFDMHQFNEFGV  
310 320 330 340 350 360

m117-1.pep 370 380 390 400 410 420  
AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTTPH  
g117-1 AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTTPH  
370 380 390 400 410 420

m117-1.pep 430 440 450 460 470 480  
GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS  
g117-1 GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS  
430 440 450 460 470 480

m117-1.pep 490 500 510 520 530 540  
VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY  
g117-1 VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY  
490 500 510 520 530 540

m117-1.pep 550 560 570 580 590 600  
KKPEDLYTAVGQGEISNRAIQACGTLNEPPVPVSETTIVKQSKIKKGGKNGVLIDGED  
g117-1 KKPEDLYTAVGQGEISNRAIQACGTLNEPPVPVSATTIVKQSKIKKGGKNGVLIDGED  
550 560 570 580 590 600

m117-1.pep 610 620 630 640 650 660  
GLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDAASWAALQEG  
g117-1 GLMTTLAKCKPAPPDDIAGFVTRERGISVHRKTCPSFRHLAEHAPEKVLDAASWAALQEG  
610 620 630 640 650 660

m117-1.pep 670 680 690 700 710 720  
QVFAVDIEIRAQDRSGLLRDVSALARHKLNVTAQTQSRDLEASMRFTELVKQVNDLPR  
g117-1 QVFAVDIEIRAQDRSGLLRDVSALARHKLNVTAQTQSRDLEASMRFTELVKQVNDLPR  
670 680 690 700 710 720

m117-1.pep 730  
VLASLGDKGVLSVTRLX  
g117-1 VLAGLGDKGVLSVTRLX  
730

m117-1/RelA

sp|P55133|RELA\_VIBSS GTP PYROPHOSPHOKINASE (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE) (PPGPP SYNTHETASE I) >gi|537617 (U13769) ppGpp synthetase I [Vibrio sp.] Length = 744  
Score = 536 bits (1366), Expect = e-151  
Identities = 288/685 (42%), Positives = 432/685 (63%), Gaps = 31/685 (4%)

Query: 74 LDLLPDAVAATLLADI---GRYVPDWNLLVSERCNSTVAELVKGVDVQKLTHFARVDSL 130  
L + D + A LL + G Y D + E + T+ LV+GV+++ ++ ++ S  
Sbjct: 68 LSMDADTLIAALLYPLVEGGCYSTD--ALKEEYSGTILHLVQGVQMCAS---QLKST 121

Query: 131 ATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTTLOFLSNAPDSPEKRAVAKETLDI 190  
A +A Q + +R+MLL+MV D R V+IKLA R L+ + + PD +RA A+E +I  
Sbjct: 122 AEETAQAAQVDNIRRMLLSMVDDFRCVVIKLAERICNLREVKDQDPDEV-RRAAAQECANI 180

Query: 191 FAPLANRLGVWQLKWQLEDLGRHQKPEKYREIALLLDEKRTTERLEYIENFLNILRGELK 250  
+APLANRLG+ QLKW++ED FR+Q P+ Y++IA L E+R +R +YI +F++ L +K  
Sbjct: 181 YAPLANRLGIGQLKWEIEDYAFRYQHPDITYKQIAQLSERRIDREDYITHFVDDLS DAMK 240

Query: 251 KYNVHFEVAGRPKHIYSIYKMKVKKLSFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQ 310  
N+ EV GRPKHIYSI++KM KK L FD LFD+RAVRI+ + + +CY LG+VH+ ++  
Sbjct: 241 ASNIRAEOVGRPKHIYSIWRKMOKKSLEFDELFDVRAVRIAEELQDCY AALGVVHTKYR 300

Query: 311 PIPGEFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRTQFDMHQFNEFGVAAHWRYKEG- 369  
+P EFDDY+ANPK NGY+S+HTV++GPE K +E+QIRT MH+ +E GVAHW+YKEG  
Sbjct: 301 HLPKEFDDYVANPKNGYQSIHTVVLGPEGKTIEIQIRTKMHHESELGVAHWRYKEGT 360

Query: 370 --GKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLT PHGKVL SLP 427  
G SAY++KI WLR+LL W+E M++SG ++ +++++F+D +Y TP G V+ LP  
Sbjct: 361 ASGGAQSAYDEKINWLRKLLAWQEEMSDSG--EMLDELRSQVFD RYVAFTPKGDVVDLP 418

Query: 428 TGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPSVNWLYE- 486  
+ ATP+DFAY +HS +G RC GAKVEG+IVP + L+ G +VEIIT KE +PS +WL  
Sbjct: 419 SNATPLDFAYHIHSEVGHRCIGAKVEGRIVPFTYHLQMGDQVEIITQKEPNPSRDWLNPN 478

Query: 487 -GWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKL--TPKPNLQELAENLGYYKP 543  
G+V S++A K+ A+ R+Q+ D G+ L+ +L K+ T K A+ K P  
Sbjct: 479 LGFVTSSRARAKVHAWFRKQDRDKNIAGKEILEAELVKIHATLKDAQYYAAKRFNVKSP 538

Query: 544 EDLYTAVGQGEIS-NRAIQKACGTLNEPPVPVSETTIVKQSKI-----KKGKNGV 594  
E+LY +G G++ N+ I +N+P + + K S+ KK ++ V  
Sbjct: 539 EELYAGIGSGDLRINQVINHINALVNKPTAEEDQQLLEKLSEASNQATSHKKPQORDAV 598

Query: 595 LIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVL D ASW 654  
+++G D LMT LA+CC+P P DDI GFVT+ RGISVHR C + L HAPE+++D W  
Sbjct: 599 VVEGVDNLMTHLARCCQIPGDDIQGFVTQGRGISVHRMDCEQLEELRHAPERI IDTVW 658

Query: 655 AALQEQGVFAVDIEIRAQDRSGLLRDVS DALARHKLNVTA VQTQ--SRDLEASMRFTLEV 712  
G + + + + A +R+GLL++++ L K+ V +++ + + M F LE+  
Sbjct: 659 GGGFVGN-YTITVRVTASERNGLLKELTNTLMNEKVKVAGMKSRVDYKKQMSIMDFE L 717

Query: 713 KQVNDLPRVLASLGDVKGVLSVTRL 737  
+ L RVL + VK V RL  
Sbjct: 718 TDLEVLGRVLKRIEQVKDVAEAKRL 742

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 421>:

all17-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA
51  ATTGCGCGAA TGGTTCGACA GCTACTGCAC CGCGCTGCCG AACAAACGATA
101 AAAAAGTTGT CTTAGCCGCC CGTTCGCTGG CGGAAGCACA TTACCCCGCC
151 GATGCCGCCA CGCCGTATGG CGAACCGCTG CCCGACCACT TCCTCGGGCG
201 GCGCAAAATG GTTCATGAAC TCGACCTGCT CCCCGATGCC GTCGCCGCCA
251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
301 TCCGAACGCT GCAACAGTAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351 AGTGCAGAAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401 AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAATGCT GCTGGCGATG
451 GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501 CCTGCAATTT TTAAGCAACG CCCCCGACAG CCCCAGAAAA CGCGCCGTCG
551 CCAAAGAAAC CCTCGACATC TTCGCCCGCG TCGCCAACCG TTTGGGCGTG
601 TGGCAGCTCA AATGGCAGCT CGAAGATTG GGCTTCGCGC ATCAAGAACC
651 CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
701 GCCTCGAATA CATCGAAAC TTCCTTAATA TCCTGCGTAC GGAACCTCAA
751 AAATACATA TCCACTTTGA AGTCGCCGCG CGTCCGAAAC ACATCTACTC
801 CATTACAAA AAAATGGTGA AGAAAAAAT CAGCTTCGAC GGGTTGTTCTG

```

```

851 ACATCCGCGC CGTGCGGATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
901 ACACCTGGGCA TTGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
951 CGACTACATC GCCAACCCGA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001 TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTGCAAT CCGCACCTTC
1051 GATATGCACC AATTCAACGA ATTCGGTGTC GCGCGGCACT GGCGTTACAA
1101 AGAGGGCGGC AAAGGCGATT CCGCCTACGA AAAAAAATC GCCTGGTTAC
1151 GCCAACTTTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCGCGC CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGCAC GGCAAAAGTCC TCTCCCTGCC CACAGGCGCG ACCCCCATCG
1301 ACTTCGCCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCGTGTGCGA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1551 ACTCGACAAA CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGTAC AAAAAGCCAG AAGACCTCTA CACCCCGGTC
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751 CCAAAATCAA AAAAGGCGGC AAAAACGGCG TGCTCATCGA CGGCGAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCGCGCG CGCCCGACGA
1851 CATTGTCGCG TTGCTTACCC GCGATCGCGG CATTTCGGTA CACCGCAAAA
1901 CCTGCCCTC TTTCCGACAC CTCGCCGAAC ACGCGCCCGA AAAAGTACTG
1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGA CAAGTGTTCC CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTTACCGA
2151 CCTCCACGC GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTA

```

This corresponds to the amino acid sequence <SEQ ID 422; ORF 117-1.a>:

a117-1.pep

```

1  MTAISPIQDT QSATLQELRE WFDSYCTALP NNDKKLVLA RSLAEAHYPA
51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLNLRTELK
251 KYNIHFEVAG RPKHIYSIYK KMKKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKNGYKS LHTVIVGPED KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401 LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQVRE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501 YIRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601 GLMTTLAKCC KPAPPDDIVG FVTRDRGISV HRKTCPSFRH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DEASMRFTL EVKQVTDLPR VLASLGDKV VLSVTRL*

```

a117-1/m117-1 97.7% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSATLQELREWFDSDYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL					
a117-1	MTAISPIQDTQSATLQELREWFDSDYCTALPNNDKKLVLAARSRLAEAHYPADAATPYGEPL					
	10	20	30	40	50	60
m117-1.pep	PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK					
a117-1	PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK					
	70	80	90	100	110	120
m117-1.pep	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK					
a117-1	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK					
	130	140	150	160	170	180
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIEN					
a117-1	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN					
	190	200	210	220	230	240
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIEN					
a117-1	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN					
	190	200	210	220	230	240



m117-1.pep	250	260	270	280	290	300
	FLNILRGELKKYNVHFEVAGRPKHIYSIYKMKVKKKLSFDGLFDIRAVRILVDTVPECYT					
a117-1	250	260	270	280	290	300
	FLNILRTELKKYNIHFEVAGRPKHIYSIYKMKVKKKLSFDGLFDIRAVRILVDTVPECYT					
m117-1.pep	310	320	330	340	350	360
	TLGIVHSLWQPIPGFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV					
a117-1	310	320	330	340	350	360
	TLGIVHSLWQPIPGFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV					
m117-1.pep	370	380	390	400	410	420
	AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH					
a117-1	370	380	390	400	410	420
	AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH					
m117-1.pep	430	440	450	460	470	480
	GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS					
a117-1	430	440	450	460	470	480
	GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS					
m117-1.pep	490	500	510	520	530	540
	VNWLYEGWVKSNAIGKIRAYIROQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY					
a117-1	490	500	510	520	530	540
	VNWLYEGWVKSNAIGKIRAYIROQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY					
m117-1.pep	550	560	570	580	590	600
	KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKKGGKNGVLIDGED					
a117-1	550	560	570	580	590	600
	KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKKGGKNGVLIDGED					
m117-1.pep	610	620	630	640	650	660
	GLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDAASWAALQEG					
a117-1	610	620	630	640	650	660
	GLMTTLAKCKPAPPDDIVGFVTRDRGISVHRKTCPSFRHLAEHAPEKVLDAASWAALQEG					
m117-1.pep	670	680	690	700	710	720
	QVFAVDIEIRAQDRSGLLRDVS DALARHKL NVTAVQTQSRDLEASMRFTLEVQVNDLPR					
a117-1	670	680	690	700	710	720
	QVFAVDIEIRAQDRSGLLRDVS DALARHKL NVTAVQTQSRDLEASMRFTLEVQVNDLPR					
m117-1.pep	730					
	VLASLGDVKGVLVSVTRLX					
a117-1	730					
	VLASLGDVKGVLVSVTRLX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 423>:

```

g118.seq
1   ATGTGCGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
51  TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAAGCTG GAGAATGATT TAATcgaGGT TAGGAGAAAA
151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
201 CATTATTGAT TTTTAAATGG TTCCAAATTG GGAGCTTTTT GAAATTAAAG
251 CTTCCTTGG GTTGCCTGAT AGCGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTGAT TATTACAaCA AAAAATAG

```

This corresponds to the amino acid sequence <SEQ ID 424; ORF 118.ng>:

```

g118.pep
1   MCEFKDFRRN IPCFEEYDEN SFIGKYYDDG VWDDEEYWKI ENDLIEVRRK
51  YPYPMIDIPRD IVIGIGTIID FLMVFNWELF EIKASPWLPD SVGIHERYER
101 FTTMLRYIFT EKDIVNVRFD YYNKK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 425>:

```
m118.seq
1   ATGTGTGAGT TCAAGGATAT TATAAGAAAC GTTCCTTATT TTGAGGGGTA
51  TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAAGTTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
151 TATCCTTATC CGATGGACAT ACCAAGATAT GTTGTCAATG GAATCGGTAC
201 CATTATTGAT TTCTTAATGG TTCCAAATG GAAACTTTT GAAATTAAAG
251 CTTCCCCTTG GTTGCCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAACA AAAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 426; ORF 118>:

```
m118.pep
1   MCEF KD IIRN VPYF EGYDEN SFIG KWYDDG VWDDEE YWKL ENDL IEVRKK
51  YPYM DIPRY VVIG IGTIID FLMV PNWKL F EIKAS PWLPD SVGI HER YER
101 FTTML RYIFT EKDI VNVRFD YYNKK*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 118 shows 92.8% identity over a 125 aa overlap with a predicted ORF (ORF 118.ng) from *N. gonorrhoeae*:

```
m118/g118

      10      20      30      40      50      60
m118.pep MCEF KD IIRN VPYF EGYDEN SFIG KWYDDG VWDDEE YWKL ENDL IEVRKK YPYM DIPRY
          |||||: ||:| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
g118     MCEF KD FRN I PCFEE YDEN SFIG KWYDDG VWDDEE YWKL ENDL IEVRKK YPYM DIPRD
          10      20      30      40      50      60

      70      80      90      100     110     120
m118.pep VVIG IGTIID FLMV PNWKL FEIKAS PWLPD SVGI HER YER FTTML RYIF TEKDI VNVRFD
          :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g118     IVIG IGTIID FLMV PNWEL FEIKAS PWLPD SVGI HER YER FTTML RYIF TEKDI VNVRFD
          70      80      90      100     110     120

m118.pep YYNKKX
          |||||
g118     YYNKKX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 427>:

a118.seq

```
1   ATGTGTGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
51  TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAATTTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
201 CATTATTGAT TTTTAAATGG TTCCAAATG GGAGCTTTT GAAATTAAAG
251 CTTCCCCTTG GTTGCCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAACA AAAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 428; ORF 118.a>:

a118.pep

```
1   MCEF KD FRN I PCFEE YDEN SFIG KWYDDG VWDDEE YWKL ENDL IEVRKK
51  YPYM DIPRD IVIG IGTIID FLMV PNWEL F EIKAS PWLPD SVGI HER YER
101 FTTML RYIFT EKDI VNVRFD YYNKK*
```

m118/a118 93.6% identity in 125 aa overlap

```
m118.pep      10      20      30      40      50      60
MCEF KD IIRN VPYF EGYDEN SFIG KWYDDG VWDDEE YWKL ENDL IEVRKK YPYM DIPRY
          |||||: ||:| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
a118         MCEF KD FRN I PCFEE YDEN SFIG KWYDDG VWDDEE YWKL ENDL IEVRKK YPYM DIPRD
          10      20      30      40      50      60
```

	70	80	90	100	110	120
m118.pep	VVIGIGTIIDFLMVPNWKLF EIKAS PWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRF					
a118	IVIGIGTIIDFLMVPNWKLF EIKAS PWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRF					
	70	80	90	100	110	120
m118.pep	YYNKXX					
a118	YYNKXX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 429>:

```

g120.seq
1   ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAAGGCTACC CCAATCCGCC GTGCTGCACT
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
201 TTTCGAATCC GCGGTACGG TTGTCGGCAA TACCCTGCAC CCTGCCTACT
251 ATAAAGACAT ACGCAGGGGC AAAGTGTATG CGGAAGCCAA ATTCGCCGAC
301 GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAACCG AGCAAAGCCC
351 CAAGGCTATG GATTTGTTCA CGCTTGCCCTG GCAGTTGGCG GCAAATGACG
401 CGAAACTCCC CCCGGGTCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451 GTCGGCGGCC TGAATAAGGC GGTACGGGA AAATACAGCA Tagggcgcggt
501 gGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATACGGTAA
551 CGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGACgaCG GCAAACCTA TACGCTGAAG CTCAAATCGG TGCAGATCAA
651 CGGACAGGCC GCCAAACCGT AA

```

This corresponds to the amino acid sequence <SEQ ID 430; ORF 120.ng>:

```

g120.pep
1   MMKTFKNIFS AAILSAALPC AYAARLPQSA VLHYSGSYGI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PAYYKDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DTVTYFFAPS LNNIPAQIGY
201 TDDGKTYTLK LKSVQINGQA AKP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 431>:

```

m120.seq
1   ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGmACT
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
201 TTTCGAGTCC GCGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT
251 ATAGAGACAT ACGCAGGGGC AAAGTGTATG CGGAAGCCAA ATTCGCCGAC
301 GGCAGCGTAA CTTACGGCAA AGCGGGCGAG AGCAAACCG AGCAAAGCCC
351 CAAGGCTATG GATTTGTTCA CGCTTGCCCTG GCAGTTGGCG GCAAATGACG
401 CGAAACTCCC CCCGGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451 GTCGGCGGTT TGAATAAGGC GGTACAGGA AAATACAGCA TAGGCGGCGT
501 GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA
551 TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGACGACG GCAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
651 CGGCCAGGCA GCCAAACCG

```

This corresponds to the amino acid sequence <SEQ ID 432; ORF 120>:

```

m120.pep
1   MMKTFKNIFS AAILSAALPC AYAAGLPQSA VLXYSYSYGI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PTYYRDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAQIGY
201 TDDGKTYTLK LKSVQINGQA AKP

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 120 shows 97.3% identity over a 223 aa overlap with a predicted ORF (ORF 120.ng) from *N. gonorrhoeae*:

m120/g120

m120.pep	10	20	30	40	50	60
	MMKTFKNIFS	AAILS	SAALPC	AYAAGLP	QSAVLX	YSGSYGIPATMTFERSGNAYKIVSTIK
g120	10	20	30	40	50	60
	MMKTFKNIFS	AAILS	SAALPC	AYAARLP	QSAVLHY	SGSYGIPATMTFERSGNAYKIVSTIK
m120.pep	70	80	90	100	110	120
	VPLYNIRFES	GGTVVGN	TLHPTTY	RDIRRG	KLYAEAK	FADGSVTYKGAGESKTEQSPKAM
g120	70	80	90	100	110	120
	VPLYNIRFES	GGTVVGN	TLHPAYY	KDIRRG	KLYAEAK	FADGSVTYKGAGESKTEQSPKAM
m120.pep	130	140	150	160	170	180
	DLFTLAWQLA	ANDAKLP	PGKITNG	GKKLYSV	GGLNKAG	TGKYSIGGVETE
g120	130	140	150	160	170	180
	DLFTLAWQLA	ANDAKLP	PGKITNG	GKKLYSV	GGLNKAG	TGKYSIGGVETE
m120.pep	190	200	210	220		
	DAVMYFFAP	SLNNIPA	QIGYTDD	GKTYTLK	LKSVQING	QAAKP
g120	190	200	210	220		
	DTVTYFFAP	SLNNIPA	QIGYTDD	GKTYTLK	LKSVQING	QAAKPX

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 433>:

a120.seq

```

1  ATGATGAAGA CTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGCACT
101  ATTCGGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
151  AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
201  TTTCGAGTCC GGCGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT
251  ATAGAGACAT ACGCAGGGGC AAAGTGTATG CGGAAGCCAA ATTCGCGGAC
301  GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAAGCCC
351  CAAGGCTATG GATTGTGTTCA CGCTTGCTTG GCAGTTGGCG GCAAATGACG
401  CGAAACTCCC CCCGGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451  GTCGGCGGTT TGAATAAGGC GGGTACAGGA AAATACAGCA TAGGCGGCGT
501  GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA
551  TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601  ACCGACGACG GCAAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
651  CGGCCAGGCA GCCAAACCGT AA

```

This corresponds to the amino acid sequence <SEQ ID 434; ORF 120.a>:

a120.pep

```

1  MMKTFKNIFS AAILSALPC AYAAGLPQSA VLHYSGSYGI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNLH PTYYRDIRRG KLYAEAKFAD
101  GSVTYKGAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151  VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAQIGY
201  TDDGKTYTLK LKSVQINGQA AKP*

```

m120/a120 99.6% identity in 223 aa overlap

m120.pep	10	20	30	40	50	60
	MMKTFKNIFS	AAILS	SAALPC	AYAAGLP	QSAVLX	YSGSYGIPATMTFERSGNAYKIVSTIK
a120	10	20	30	40	50	60
	MMKTFKNIFS	AAILS	SAALPC	AYAAGLP	QSAVLHY	SGSYGIPATMTFERSGNAYKIVSTIK

343

	10	20	30	40	50	60
m120.pep	70	80	90	100	110	120
	VPLYNIRFESGGTVVGNLHPTYYRDIRRGKLYAEAKFADGSVITYGKAGESKTEQSPKAM					
a120	VPLYNIRFESGGTVVGNLHPTYYRDIRRGKLYAEAKFADGSVITYGKAGESKTEQSPKAM					
	70	80	90	100	110	120
m120.pep	130	140	150	160	170	180
	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEVEVKYRVRRGD					
a120	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEVEVKYRVRRGD					
	130	140	150	160	170	180
m120.pep	190	200	210	220		
	DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX					
a120	DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX					
	190	200	210	220		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 435>:

g121.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
51  GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCTGACC GGTTCGCGCG CAAATGCTG
151 GATTTGCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATAACAG TTGCCGATTT
351 GCCGCTGCTG GCGGAAGTga cgcggatttt TACCGTCggc gacttcCGCA
401 GCCGCGACCT TGCTGCCGCG GGacaAGGTG CGCCGCTCGT CCCCgcCTTT
451 CACGAAGCCC TGTTCCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCggCGCA CCGCCTTCG
551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
601 cactTGGcagc TGCCTTACGA CAAAaacggt gcAAAGgcgg cacAAGGCAA
651 catatTGcCg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTCTCTAC
701 AACCccaccc aaAAAGCACG GGgcGCGaac TgtttgcccT AAattggtc
751 gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgct
801 ttccccgattc accgcgcaaA ccgTttggga cgccgtctca CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGCGCG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCTGCGACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTtg
1001 cgtggttggC GCGGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCGACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 436; ORF 121.ng>:

g121.pep

```

1  METQLYIGIM SGTSMDGADA VLVRMDGGKW LGAEGHAFTP YPDRLRKRL
51  DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
101 TVRHAPEHGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPFA
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTPGP NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSEF TAQTVWDVAH HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 437>:

m121.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CCAATGCTG

```

```

151 GATTTCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTG
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCACCGTCC GACATTACCG CCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATT
351 GCCGCTGCTG GCGxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
401 xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
451 xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
501 xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
551 xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
601 xxxxxxCAGC TTCCTTACGA CAAAAACGGT GCAAAGTCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTCGCAC
701 AACGCCACCC TAAAAGCAGC GGGCGCGAAC TGTTTGCCAT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGTTT ACCGCGCAA CCGTTTGC GAACGATCCG CAATCCTGTT
851 CAGATGCCCG TCAAATGTAC ATTTGCGACG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGAAGCC GCCGnATTG
1001 CGTGGTTGGC GGCCTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 438; ORF 121>:

m121.pep

```

1 METQLYIGIM SGTSMGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRQLL
51 DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHGYSIQLADLPLL Axxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
151 xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx xxxxxxxxxxxx
201 xxQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251 ETYLDGGENR YDVLRTLRSR TAQTVCDVAVS HAAADAROMY ICDGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis* *menA* with *menB*

ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from *N. gonorrhoeae*:

m121/g121

```

          10      20      30      40      50      60
m121.pep  METQLYIGIMSGTSMGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
          |||
g121       METQLYIGIMSGTSMGADAVLVRMDGGKWLGAEGHAFTPYPDRLRRKLLDLQDTGTDEL
          10      20      30      40      50      60
          70      80      90     100     110     120
m121.pep  HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
          |||
g121       HRSRMLSQELSRLYAQTAAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
          70      80      90     100     110     120
          130     140     150     160     170     180
m121.pep  AXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
          | : : : : :
g121       AELTRIFTVGDFRSRDLAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
          130     140     150     160     170     180
          190     200     210     220     230     240
m121.pep  XXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST
          : : : |||
g121       PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
          190     200     210     220     230     240
          250     260     270     280     290     300
m121.pep  GRELFAINWLETYLDGGENRYDVLRTLRSRFTAQTVCDVAVSHAAADAROMYICDGGIRNPV
          |||
g121       GRELFALNWLETYLDGGENRYDVLRTLRSRFTAQTVWDVAVSHAAADAROMYICGGGIRNPV
          250     260     270     280     290     300

```

345

	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVSLHSTADLNLDPOWVEAAAFWLAACWINRIPGSPHKATGASKPCIL					
	:					
g121	LMADLAECFGTRVSLHSTAE LNLDPOWVEAAAFWLAACWINRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYY					
g121	GAGYYYY					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 439>:

```

a121.seq
1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCGCGCA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCACGCA CCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCTTACGA CAAAACGGT GCAAAGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACCCACCC TAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGATT ACCGCGCAA CCGTTTTCGA CGCGTCTCA CACGACGGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 440; ORF 121.a>:

```

a121.pep
1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPESY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAFLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPGP NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

m121/a121 74.0% identity in 366 aa overlap

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
	:					
a121	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
	10	20	30	40	50	60
m121.pep	HRSRILSQELSRLYAQTAAELLCSONLAPSDITALGCHGQTVRHAPESHSYQLADLPLL					
	:					
a121	HRSRILSQELSRLYAQTAAELLCSONLAPSDITALGCHGQTVRHAPESHSYQLADLPLL					
	70	80	90	100	110	120

	130	140	150	160	170	180
m121.pep	AXXX					
a121	AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST					
a121	PAFGFDTGPGNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICDGGIRNPV					
a121	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFVDAVSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGTRVSLHSTAE LNLDPQWVEAAAFWMAACWVNRIIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYY					
a121	GAGYYYY					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 441>:

m121-1.seq

```

1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCAGCG CTTTACCCCC TACCCCGGCA GGTACGCCG CCAATTGCTG
151 GATTGTCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAACCTC CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATAACGC TTGCCGATT
351 GCCGCTGCTG GCGGAACGGA CGCGGATTTT TACCGTCGGC GACTTCCGCA
401 GCCCGACCTT TCGCGCCGGC GGACAAGGCG CGCCACTCGT CCCGCTT
451 CACGAAGCCC TGTTCCGCGA CAACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCGACGCA CCCGCTTCG
551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGACGCGTG GACGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACCCACCCC TAAAGCACG GGGCGCGAAC TGTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TCGGGACGCT
801 TTCCCGTTTT ACCGCGCAAA CCGTTTGC GAACGCTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGNATTG
1001 CGTGGTTGGC GCGGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG ANCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 442; ORF 121-1>:

m121-1.pep

```

1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGLRRQLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPEHGY SIQLADLP LL AERTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDNRE TRAVLNIGGI ANISVLPPDA PAFGFDTPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVCDVSH AAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYYY*

```

m121-1/g121 95.6% identity in 366 aa overlap

10 20 30 40 50 60



```

m121-1.pep  METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
g121         METQLYIGIMSGTSMGDADAVLVRMDGGKWLGAEGHAFTPYPDRLRRKLLDLQDTGTDEL
              10      20      30      40      50      60

              70      80      90      100     110     120
m121-1.pep  HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
g121         HRSRMLSQELSRLYAQTAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
              70      80      90      100     110     120

              130     140     150     160     170     180
m121-1.pep  AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA
g121         AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
              130     140     150     160     170     180

              190     200     210     220     230     240
m121-1.pep  PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
g121         PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
              190     200     210     220     230     240

              250     260     270     280     290     300
m121-1.pep  GRELFALNWLETYLDGGENRYDVLRLTSRFTAQTVCDVASHAAADARQMYICGGGIRNPV
g121         GRELFALNWLETYLDGGENRYDVLRLTSRFTAQTVWDVASHAAADARQMYICGGGIRNPV
              250     260     270     280     290     300

              310     320     330     340     350     360
m121-1.pep  LMADLAECFGTRVSLHSTADLNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
g121         LMADLAECFGTRVSLHSTAE LNLDPOWVEAAAFWLAACWINRIPGSPHKATGASKPCIL
              310     320     330     340     350     360

m121-1.pep  XAGYYYYX
g121         GAGYYYYX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 443>:

```

a121-1.seq
1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51 GCGGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCAGCG CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAAC TGACGAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCTGTG ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAACCTC CGCGCGCTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCC GCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCAGCGCA CCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCCGAC
701 AACCCACCCC TAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGATTG ACCGCGCAAA CCGTTTTCGA CGCCGCTCTC CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 444; ORF 121-1.a>:

```

a121-1.pep
1  METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL
51 DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCQNLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA

```

201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL  
 251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV  
 301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK  
 351 ATGASKPCIL GAGYYY\*

m121-1/a121-1 96.4% identity in 366 aa overlap

m121-1.pep	10	20	30	40	50	60
	METQLYIGIMSGTSM	GADAVLIRMDGGKWL	GAEGHAFTPY	PGRRLRQLLDL	QDTGADEL	
a121-1	METQLYIGIMSGTSM	GADAVLIRMDGGKWL	GAEGHAFTPY	PGRRLRQLLDL	QDTGADEL	
	10	20	30	40	50	60
m121-1.pep	70	80	90	100	110	120
	HRSRILSQELSR	LYAQTAELLCSQ	NLAPSDITALG	CHGQTVRHA	PEHGYISQL	ADLPL
a121-1	HRSRILSQELSR	LYAQTAELLCSQ	NLAPSDITALG	CHGQTVRHA	PEHGYISQL	ADLPL
	70	80	90	100	110	120
m121-1.pep	130	140	150	160	170	180
	AERTRIFTVGDF	RSRDLAAGGQGA	PLVPAPFHEAL	FRDNRETRAV	LNIGGIANIS	VLPPDA
a121-1	AERTRIFTVGDF	RSRDLAAGGQGA	PLVPAPFHEAL	FRDNRETRAV	LNIGGIANIS	VLPPDA
	130	140	150	160	170	180
m121-1.pep	190	200	210	220	230	240
	PAFGFDTGPGN	MLMDAWTQAHW	QLPYDKNGAKA	AAQGNILPQL	LDRLLAHPY	FAQPHPKST
a121-1	PAFGFDTGPGN	MLMDAWTQAHW	QLPYDKNGAKA	AAQGNILPQL	LDRLLAHPY	FAQPHPKST
	190	200	210	220	230	240
m121-1.pep	250	260	270	280	290	300
	GRELFALNWLE	TYLDGGENRYD	VLRTLSRFTA	QTVCDVSHA	AADARQMYI	CGGGIRNPV
a121-1	GRELFALNWLE	TYLDGGENRYD	VLRTLSRFTA	QTVCDVSHA	AADARQMYI	CGGGIRNPV
	250	260	270	280	290	300
m121-1.pep	310	320	330	340	350	360
	LMADLAECFG	TRVSLHSTAD	LNLDPQWVEA	AAXFAWLA	ACWINRIPG	SPHKATGASKPCIL
a121	LMADLAECFG	TRVSLHSTAD	LNLDPQWVEA	AAXFAWLA	ACWINRIPG	SPHKATGASKPCIL
	310	320	330	340	350	360
m121-1.pep	XAGYYYYX					
a121	GAGYYYYX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 445>:

g122.seq

```

1  ATGGCTTTAC TGAGCATCCG CAAGCTGCAC AAACAATACG GCAGCGTAAC
51  CGCCATCCAA TCCTTAGACT TGGACTTGGA AAAAGGCGAA GtcatCGTAC
101 TGCTGGGCCC gTccggtcgc ggCAAATCCA CCTCctgcg ctgcgtaac
151 GGTTTGGAGC CGCACCAagg cgGCAGCATC GTGATGGACG GTgtcgGCGA
201 ATTCggcAAA GACGTTTCCT GGCAAACCGC CCGGCAAAaA gtcggtatgg
251 tctttcaaag taacgAactg Tttgcccaca tgaccgtcat cgAaaacatc
301 ttcttAggcC CGGTAAagga aCAAAACgc gaccgtgccg aagcaGAGGC
351 gCAAGCCGGC AAactGttgg aacgcgTCGG actgctAGAC CGCAAAACG
401 CCTATCCGCG CGAACTTTCC GGCGGTCAGA AACAGCGCAT CGCCATTGTC
451 CGCGCCCTGT GCCTGAATCC GGAAGTCATC CTGCTGGACG AAATCACCGC
501 CGCACTTGAC CCCGAAATGG TGC CGAAGT CTTGGAAGTG GTTTTGGAAC
551 TCGCCCGCGA AGGGATGAGT ATGCTCATCG TAACCCACGA AATGGGGTTC
601 GCACGCAAAG TTGCCGACCG CATCGTCTTT ATGGACAAAG GCGCATCGT
651 CGAATCGTCC GACCCGAAA CCTTTTTTTC CGCACC AAAA AGCGAACCGC
701 CCCGCCAATT TCTGGCAGGT ATGGACTACT GA
```

This corresponds to the amino acid sequence <SEQ ID 446; ORF 122.ng>:

1	MALLSIRKLM	KQYGSVTAIQ	SLDLDLEKGE	VIVLLGPSGC	GKSTLLRCVN
51	GLEPHQGGSI	VMDGVGEFGK	DVSWQTARQK	VGMVFQSNEL	FAHMTVIENI
101	FLGPVKEQNR	DRAEAEAQAG	KLLERVGLLD	RKNAYPRELS	GGQKQRIAIV
151	RALCLNPEVI	LDEITAAALD	PEMVEVLELV	VELAREGMS	MLIVTHEMGF
201	ARKVADRIIV	MDKGGIVESS	DPETFFSAPK	SERAROFLEG	MDY*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 447>:

m122.seq

1	GTTGTCATGA	TTAAAAATCCG	CAATATCCAT	AAGACCTTTG	GCGAAAAACAC
51	TATTTTGC	GGCATCGATT	TGGATGTGTG	CAAAGGGCAG	GTGGTCGTCA
101	TCCTCGGGcC	TTCCGGCTCA	GGCAAAACGA	CGTTTTCTGC	ATGCCTAAAC
151	GCGTTGGAAA	TGCCCGAAGA	CGGACAAATC	GAGTTGCAGA	ACGAGCGACC
201	GCTGAAAATC	GATTTTCTA	AAAAACCAAG	CAAAACACGAT	ATTTTGGCAC
251	TGCGCCGCAA	ATCAKGCATG	GTGTTTCAAC	AATACAAyCT	CTTTCCGCAC
301	AAAACCGCCT	TGGAACACGT	AATGGAAGGA	CCGGTTGCCG	TACAgGGCAA
351	GCCTGCCGCC	CAAGCGCGCG	AAGAGGCTCT	GAAACTGCTG	GAAAAAGTCG
401	GCTTGGGCGA	CAAAGTGGAT	TTGTATCCCT	ACCAGCTTTT	CGGCGGTCAG
451	CAGCAGCGCG	TCGGCATTGC	CCGCGCATTG	GCGATTACAG	CTGAACTGAT
501	GCTGTTTGAC	GAACCGACTT	CCGCGCTCGA	TCCTGAATTG	GTGCAAGATG
551	TTTTGGATmC	CATGAAGGAA	TTGGCGCAAG	AAGGCTGGAC	CATGGTTGTC
601	GTTACGCATG	AAATCAAGTT	CGCCTTAGAA	GTGGCAACCA	CCGwCGTCGT
651	GATGGACrC	GGCGTTATTG	TCGAACAAGG	CAGCCCGCAA	GATTTGTTTCG
701	ACCACCCCAA	ACCGAACGG	ACGCGGAGAT	TTTTAAGCCA	AATCCAATCT
751	ACCAAGATTT	GA			

This corresponds to the amino acid sequence <SEQ ID 448; ORF 122>:

m122.pcp

1	VVMIKIRNIH	KTFGENTILR	GIDL DVCKGQ	VVVLGP SG S	GKTTFLRCLN
51	ALEMPEDGQI	EFDNERPLKI	DFSKKPSKHD	ILALRRKSXM	VFQQYNLFPH
101	KTALENVMEG	PVAVQGKPA A	QAREEALKLL	EKVGLGDKVD	LYPYQLSGGQ
151	VQRVGIARAL	AIQPELM LFD	EPTSALDP EL	VQDVLDXMKE	LAQEGWTMVV
201	VTHEIKFALE	VATTXVVM DX	GVIVEQGS PQ	DLFDHPKHER	TRRFLS QIQS
251	TKI *				

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 122 shows 47.2% identity over a 246 aa overlap with a predicted ORF (ORF 122.ng) from *N. gonorrhoeae*:

m122/q122

BNSDOCID: <WO 9957280A2 | >

350

	180	190	200	210	220	230
	250					
m122.pep	TRRFLSQIQSTKIX					
	: : : :					
g122	ARQFLAGMDYX					
	240					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 449>:

```

a122.seq
1  GTTGT CATGA  TTAAATCCG  CAATATCCAT  AAGACCTTCG  GCAAAAATAC
51  CATTTTGCGC  GGCATCAATT  TGGATGTGTG  CAAAGGGCAG  GTGGTCGTCA
101 TCCTCGGGCC  TTCCGGCTCA  GGCAAAACGA  CGTTTCTGCG  ATGCCTAAAC
151 GCGTTGGAAG  TGCCCGAAGA  CGGACAAATC  GAGTTCGACA  ACAGCGGACC
201 GCTGAAAATC  GATTTTTTCTA  AAAAACCAG  CAAACACGAT  ATTTTGGCAC
251 TGCGCCGCAA  ATCAGGCATG  GTGTTTCAAC  AATAACAACCT  CTTTCCGCAC
301 AAAACCGCCT  TGGAAAACGT  GATGGAAGGA  CCGGTTGCCG  TACAGGGCAA
351 GCCTGCCGCC  CAAGCGCGCG  AAGAGGCTCT  GAAACTGCTG  GAAAAAGTCG
401 GCTTGGGCGA  CAAAGTGGAT  TTGTATCCCT  ACCAGCTTTC  CGGCGGTCAG
451 CAGCAGCGCG  TCGGCATTGC  CCGAGCATTG  GCGATTACAG  CCGAGCTGAT
501 GTTGTGTTGAC  GAACCCACTT  CCGCGCTTGA  CCCCAGATTG  GTGCAAGACG
551 TGTGAACGC  CATGAAGGAA  TTGGCGCGGG  AAGGTTGGAC  GATGGTCGTC
601 GTTACCCACG  AAATCAAGTT  CGCGCTGGAA  GTTGCCACGA  CCGTTGTCGT
651 GATGGACGGC  GCGGTTATCG  TAGAGCAGGG  CAGCCCGAAA  GAGTTGTTCC
701 ACCACCCCAA  ACACGAACGG  ACGCGGAGAT  TTTTAAGCCA  AATCCAATCT
751 ACCAAGATT  GA

```

This corresponds to the amino acid sequence <SEQ ID 450; ORF 122.a>:

```

a122.pep
1  VVMIKIRNIH  KTFGKNTILR  GINLDVCKGQ  VVILGPSGS  GKTTFRLCLN
51  ALEMPEDGQI  EFDNERPLKI  DFSKKPSKHD  ILALRRKSGM  VFQQYNLFPH
101 KTALENVMEG  PVAVQGKPAA  QAREEALKLL  EKVGLGDKVD  LYPYQLSGGQ
151 QQRVGIARAL  AIQPELMLFD  EPTSALDPEL  VQDVLNAMKE  LAREGWTMVV
201 VTHEIKFALE  VATTVVVMDG  GVIVEQGSFK  ELFDHPKHER  TRRFLSQIQS
251 TKI*

```

m122/a122 96.0% identity in 253 aa overlap

	10	20	30	40	50	60
m122.pep	VVMIKIRNIHKTFGENTILRGIDLVDVCKGQVVILGPSGSGKTTFLRCLNALEMPEDGQI					
a122	VVMIKIRNIHKTFGKNTILRGINLDVCKGQVVILGPSGSGKTTFLRCLNALEMPEDGQI					
	10	20	30	40	50	60
	70	80	90	100	110	120
m122.pep	EFDNERPLKIDFSKKPSKHDILALRRKSXMFVQQYNLFPHKTALENVMEGPPVAVQGKPAA					
a122	EFDNERPLKIDFSKKPSKHDILALRRKSGMVFQQYNLFPHKTALENVMEGPPVAVQGKPAA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m122.pep	QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQQRVGIARALAIQPELMLFDEPTSALDPEL					
a122	QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQQRVGIARALAIQPELMLFDEPTSALDPEL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m122.pep	VQDVLDMKELAQEGWTMVVVTHEIKFALEVATTXVMDXGVIVEQGSPODLFDHPKHER					
a122	VQDVLNAMKELAREGWTMVVVTHEIKFALEVATTVVMDGGVIVEQGSPELFDHPKHER					
	190	200	210	220	230	240
	250					
m122.pep	TRRFLSQIQSTKIX					

a122  
| | | | | | | | | |  
TRRFLSQIQSTKIX  
250

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 451>:

g122-1.seq

```
1  ATGATTAAAA  TCCGCAATAT  CCATAAGACC  TTTGGCGAAA  ACACCATTTT
51  GCGCGGCATC  GATTTGATG  TGGGCAAAGG  GCAGGTGGTC  GTCATCCTCG
101 GGCCTTCCGG  CTCGGGTAAA  ACAACATTTT  TGCCTGCCT  AAACGCGTTG
151 GAAATGCCCG  AAGACGGACA  AATCGAGTTC  GACAACGCGC  GGCCGTTACG
201 CATTGATTTT  TCCAAAAAAA  CAAGCAAACA  CGATATTTTG  GCACTGCGCC
251 GCAAGTCCGG  AATGGTATTC  CAACAATACA  ACCTCTTCCC  GCATAAAACC
301 GTGTTGAAA  ACGTGATGGA  AGGGCCGGTT  GCCGTACAGG  GCAAGCCTGC
351 CGCCCAAGCG  CGCGAAGAGG  CTTTGAAACT  GCTGGA AAAA  GTCGGCTTGG
401 GCGATAAAGT  GGATTGTAT  CCCTACCAGC  TTTCCGGCGG  TCAGCAGCAG
451 CGTGTCGGTA  TCGCCCGCGC  ACTGGCGATT  CAGCCTGAAT  TGATGCTGTT
501 TGACGAACCC  ACTTCCGCGC  TGGACCCCGA  GTTGGTGCAA  GACGTGTTGG
551 ACGCCATGAA  GGAATTGGCG  CGGGAAGGTT  GGACGATGGT  CGTCGTTACC
601 CACGAAATCA  AGTTCACGCT  GGAAGTTGCC  ACGAACGTCG  TCGTGATGGA
651 CGCGGGCGTT  ATCGTAGAGC  AGGCAGCCCC  GAAAGAGTTG  TTCGACCACC
701 TCAAACACGA  ACGGACGCGG  AGATTTTTAA  GCCAAATCCA  ATCTGCCAAG
751 ATTTGA
```

This corresponds to the amino acid sequence <SEQ ID 452; ORF 122-1.ng>:

g122-1.pep

```
1  MIKIRNIHKT  FGENTILRGI  DLDVKGQV  VILGPSGSGK  TTFLRCLNAL
51  EMPEDGQIEF  DNRPLRIDF  SKKTSKHDIL  ALRRKSGMVF  QQYNLFPHKT
101 VLENVMEGPV  AVQGKPAAQA  REEALKLLEK  VGLGDKVDLY  PYQLSGGQQQ
151 RVGIARALAI  QPELMLFDEP  TSALDPELVQ  DVLDMKELA  REGWTMVVVT
201 HEIKFTLEVA  TNVVMDGGV  IVEQGSPEL  FDHLKHERTR  RFLSQIQSAK
251 I*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 453>:

m122-1.seq

```
1  ATGATTAAAA  TCCGCAATAT  CCATAAGACC  TTTGGCGAAA  AACTATTTT
51  GCGCGGCATC  GATTTGATG  TGTGCAAAGG  GCAGGTGGTC  GTCATCCTCG
101 GGCCTTCCGG  CTCAGGCAAA  ACGACGTTTC  TCGCATGCCT  AAACGCGTTG
151 GAAATGCCCG  AAGACGGACA  AATCGAGTTC  GACAACGAGC  GACCGCTGAA
201 AATCGATTTT  TCTAAAAAAC  CAAGCAAACA  CGATATTTTG  GCACTGCGCC
251 GCAAAATCAGG  CATGGTGTTT  CAACAATACA  ACCTCTTCCC  GCACAAAACC
301 GCCTTGAAA  ACGTAATGGA  AGGACCGGTT  GCCGTACAGG  GCAAGCCTGC
351 CGCCCAAGCG  CGCGAAGAGG  CTCTGAAACT  GCTGGA AAAA  GTCGGCTTGG
401 GCGCAAAAGT  GGATTGTAT  CCCTACCAGC  TTTCCGGCGG  TCAGCAGCAG
451 CGCGTCGGCA  TTGCCCGCGC  ATTGGCGATT  CAGCCTGAAC  TGATGCTGTT
501 TGACGAACCG  ACTTCCGCGC  TCGATCCTGA  ATTGGTGCAA  GATGTTTGG
551 ATACCATGAA  GGAATTGGCG  CAAGAAGGCT  GGACCATGGT  TGTCGTTACG
601 CATGAAATCA  AGTTCGCCCT  AGAAGTGGCA  ACCACCGTCG  TCGTGATGGA
651 CGCGGGCGTT  ATTGTGCAAC  AAGGCAGCCC  GCAAGATTTG  TTCGACCACC
701 CCAAACACGA  ACGGACGCGG  AGATTTTTAA  GCCAAATCCA  ATCTACCAAG
751 ATTTGA
```

This corresponds to the amino acid sequence <SEQ ID 454; ORF 122-1>:

m122-1.pep

```
1  MIKIRNIHKT  FGENTILRGI  DLDVCKGQV  VILGPSGSGK  TTFLRCLNAL
51  EMPEDGQIEF  DNERPLKIDF  SKKPSKHDIL  ALRRKSGMVF  QQYNLFPHKT
101 ALENVMEGPV  AVQGKPAAQA  REEALKLLEK  VGLGDKVDLY  PYQLSGGQQQ
151 RVGIARALAI  QPELMLFDEP  TSALDPELVQ  DVLDTMKELA  QEGWTMVVVT
201 HEIKFALEVA  TTVVMDGGV  IVEQGSPQDL  FDHPKHERTR  RFLSQIQSTK
251 I*
```

m122-1/g122-1 94.8% identity in 251 aa overlap

	10	20	30	40	50	60
m122-1.pep	MIKIRNIHKT	FGENTILRG	IDLDVCKG	QVVVILGP	SGSGKTTF	FLRCLNA
g122-1	MIKIRNIHKT	FGENTILRG	IDLDVCKG	QVVVILGP	SGSGKTTF	FLRCLNA

352

	10	20	30	40	50	60
m122-1.pep	70	80	90	100	110	120
	DNERPLKIDFSKKPSKHDILALRRKSGMVFQQYNLFPHKTALENVMEGPPVAVQGKPAAQA					
g122-1	70	80	90	100	110	120
	DNARPLRIDFSKKPSKHDILALRRKSGMVFQQYNLFPHKTVLENVMEGPPVAVQGKPAAQA					
m122-1.pep	130	140	150	160	170	180
	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					
g122-1	130	140	150	160	170	180
	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					
m122-1.pep	190	200	210	220	230	240
	DVLDTMKELAQEGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGSPODLFDHPKHERTR					
g122-1	190	200	210	220	230	240
	DVLDMKELAREGWTMVVVTHEIKFTLEVATNVVVMDGGVIVEQGSPELFDHLKHERTR					
m122-1.pep	250					
	RFLSQIQSTKIX					
g122-1	250					
	RFLSQIQSAKIX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 455>:

## a122-1.seq

```

1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTCGGCAAAA ATACCATTTT
51  GCGCGGCATC AATTGCGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCTTCCGG CTCAGGCAAA ACGACGTTTC TGCGATGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAATCAGG CATGGTGTTC CAACAATACA ACCTCTTTCC GCACAAAACC
301 GCCTTGGAAG ACGTGATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAGAA GTCGGCTTGG
401 GCGACAAAGT GGATTTGTAT CCCTACCAGC TTCCGCGCG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCGAGC ATTGGCGGAT CAGCCCGAGC TGATGTTGTT
501 TGACGAACCC ACTTCCGCGC TTGACCCCGA GTGGGTGCAA GACGTGTTGA
551 ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCGCGCT GGAAGTTGCC ACGACCGTTG TCGTGATGGA
651 CGCGCGCGTT ATCGTAGAGC AGGCGAGCCC GAAAGAGTTG TTCGACCACC
701 CCAAACACGA ACGGACGCGG AGATTTTAA GCCAAATCCA ATCTACCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 456; ORF 122-1.a>:

## a122-1.pep

```

1  MIKIRNIHKT FGKNTILRGI NLDVCKGQVV VILGPSGSGK TFLRCLNAL
51  EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQGKPAAQA REEALKLLEK VLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLNAMKELA REGWTMVVVT
201 HEIKFALEVA TTVVVMDGGV IVEQGSPEL FDHPKHERTR RFLSQIQSTK
251 I*

```

a122-1/m122-1 97.2% identity in 251 aa overlap

	10	20	30	40	50	60
a122-1.pep	MIKIRNIHKTFGKNTILRGINLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					
m122-1	MIKIRNIHKTFGENTILRGIDLVDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					
	10	20	30	40	50	60
a122-1.pep	70	80	90	100	110	120
	DNERPLKIDFSKKPSKHDILALRRKSGMVFQQYNLFPHKTALENVMEGPPVAVQGKPAAQA					
m122-1	70	80	90	100	110	120
	DNERPLKIDFSKKPSKHDILALRRKSGMVFQQYNLFPHKTALENVMEGPPVAVQGKPAAQA					
a122-1.pep	130	140	150	160	170	180
	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					
m122-1	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					

		130	140	150	160	170	180
		190	200	210	220	230	240
a122-1.pep		DVLNAMKELAREGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGGSPKELFDHPKHERTR					
		::					
m122-1		DVLDTMKELAQEGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGGSPQDLFDHPKHERTR					
		190	200	210	220	230	240
		250					
a122-1.pep		RFLSQIQSTKIX					
m122-1		RFLSQIQSTKIX					
		250					

७३  
७४

```

g125.seq
1  ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCGCCGCCA TCGGGCTGGT
51  TTGGTTCCGC GCGGCGGTAT CGATTGCCGA AATCAGCACG GGTACGCTGC
101 TCGCCCCCTT GGGCTGGCAG CCGGGTCTGG CGGGCTGTCT TTTGGGTCA
151 GCCGTCCGGC GCGCGCTGTT TTTGCGGGC GCGTATATCG CGCAGCTGAC
201 CGGACGCAGC TCGATGGAAA GTGTGCGCCT GTCGTTCCGC AAATCGGGTT
251 CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
301 GTGATGATTT ACGTCGGCGC AacggTCAGC TCCGCTTTGG GCAAAGTGTT
351 GTGGGACggc gaATCCTTTG TCTGGTGGGG ATTGGCAAA C GCGCAGTGA
401 TCGTGCTGTG GCTGGTTTTC GGCACACGCA GAACGGGCGG GCTGAAATAC
451 GTTTCGATGC TGCTGATGCT GCTTGCCGTG TGTGGTTTGA GCGTCGAAGT
501 GTTCGCTTCG TCCGGCACAA ACGCCGCGCC CGCCGTTTCA GACGGCATGA
551 CCTTCGGAAC GGCAGTCGAA CTGTCCGCCG TCATGCCGCT TTCCTGGCTG
601 CCGCTGGCCG CCGACTACAC GCGCCAAGCA CGCCGCCCGT TCGCGCAAC
651 CCGTACGGCA ACGCTCGCCT ATACGCTGAC GGGCTGCTGG ATGTATGCTT
701 TGGGTTTGGC GCGCGCTCTG TTTACCGGAG AAACCGACGT GCGGAAAATC
751 CTGTTGGGCG CGGGCTTGGG CATAACGGGC ATTCTGGCAG TGCTCTCTC
801 CACCGTTACC ACAACGTTTC TCGATACCTA TTCCGCCGGC GCGAGTGCAG
851 ACACATTTT CGCGCGTTT GCGGAAATAC CCGTCGCTGT CGCGTTACC
901 CTGatccgca ccgtgcttgc cgtcatgtg cccgttaccg aatataaaaa
951 cttctcgtgc cttatccgct cggtatcttg gccgatggcg ggtggttttg
1001 attgccqaCT TTTttgtctt AAAACGGCGT GA

```

This corresponds to the amino acid sequence <SEQ ID 458; ORF 125.ng>:

७३  
७४

m125.seq					
1	ATGTCGGGCA	ATGCCTCCTC	TCCTTTCATCT	TCCTCCGCCA	TGGGCTGAT
51	TTGGTTCGGC	GCGGCGGTAT	CGATTGCCGA	AATCAGCACG	GGTACGCTGC
101	TTGCGCCTTT	GGGCTGGCAG	CGCGGTCTCG	CGGCTCTACT	TTTGGGTCA
151	GCCGTGGCG	GCGCGTGTT	TTTTGCGGG	GCGTATATCG	CGCATCTGAC
201	CGGACGCAGC	TCGATGAAA	CGGTGCGCCT	TCGTTCGGC	AAACGCGGTT
251	CAGTGCTGTT	TTCCGTGGCG	AATATGCTGC	AACTGGCCGG	CTGGACGGCG
301	GTGATGATTT	ACGCCGCGC	AACGGTCAGC	TCCGCTTTGG	GCAAAGTGTT
351	GTGGGACGGC	GAATCTTTTG	TCTGGTGGGC	ATTGGCAAAC	GGCGCGCTGA
401	TTGTGCTGTG	GCTGGTTTTCT	GGCGCACGCA	AAACAGCCGG	GCTGAAAACC
451	GTTTCGATGC	TGCTGATGCT	GTTGGCGGTT	CTGTGGCTGA	GTCGCCAAGT
501	CTTTCCACG	GCAGGCAGCA	CCGCCGCACA	GGTTTCAGAC	GGCATGAGTT
551	TCGGAACGGC	AGTCGAGCTG	TCCGCCGTGA	TGCCGCTTTC	CTGGCTGCCG
601	CTTGCCGCCG	ACTACACGCG	CCACGCGCGC	CGCCCGTTTG	CGGCAACCCT
651	GACGGCAACG	CTCGCCTACA	CGCTGACCGG	CTGCTGGATG	TATGCCTTGG
701	GTTTGGCAGC	GGCGTTGTTC	ACCGGAGAAA	CCGACGTGGC	AAAAATCTCT
751	CTGGGCGCAr	GTTTGgGTGC	GGCAGGCATT	TTGGCGGTGC	TCCTCTCCAG
801	CGTTACCACA	ACGTTTCTCG	ATGCCTATT	CGCCGGCGCG	AGTGCGAAC

851 ACATTTCCGC GCGTTTTGCG GAAACACCCG TCGCTGTCrG CGTTACCCTG  
 901 ATCGGCACGG TACTTGCCGT CATGCTGCCC GTTACCGAAT ATGAAAACCTT  
 951 CCTGCTGCTT ATCGGCTCGG TATTTGCGCC GATGGCGGgC GGTTTTGATT  
 1001 GCCGACTTTT TCGTCTTGAA ACGGCGTGA

This corresponds to the amino acid sequence <SEQ ID 460; ORF 125>:

m125.pep  
 1 MSGNASSPSS SSAIGLIWFG AAVSIAEIST GTLLAPLGWQ RGLAALLLGH  
 51 AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFSVA NMLQLAGWTA  
 101 VMIYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT  
 151 VSMLLMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP  
 201 LAADYTRHAR RPFATLTAT LAYTLTGCM YALGLAAALF TGETDVAKIL  
 251 LGAXLGAAGI LAVVLSTVTT TFLDAYSAGA SANNISARFA ETPVAVXVTL  
 301 IGTVLAVMLP VTEYENFLLL IGSVFAPMAG GFDCRLFRLE TA\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 125 shows 92.1% identity over a 343 aa overlap with a predicted ORF (ORF 125.ng) from *N. gonorrhoeae*:

m125/g125

m125.pep	10	20	30	40	50	60
	MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQORGLAALLLGHAVGGALFFAA					
g125	MSGNASSPSSSSAIGLVWFGAAVSIAEISTGTLLAPLGWQORGLAALLLGHAVGGALFFAA					
	10	20	30	40	50	60
m125.pep	70	80	90	100	110	120
	AYIGALTGRSSMESVRLSFGKRGSVLFSVANMLQLAGWTAVMIYAGATVSSALGKVLWDG					
g125	AYIGALTGRSSMESVRLSFGKCGSVLFSVANMLQLAGWTAVMIYVAGATVSSALGKVLWDG					
	70	80	90	100	110	120
m125.pep	130	140	150	160	170	179
	ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQ-VS					
g125	ESFVWWALANGALIVLWLVFGARRTGGLKTVSMLLMLLAVLWLSVEVFASSGTNAAPAVS					
	130	140	150	160	170	180
m125.pep	180	190	200	210	220	239
	DGMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCMYALGLAAAL					
g125	DGMTFGTAVELSAVMPLSWLPLAADYTRQARRPFAATLTATLAYTLTGCMYALGLAAAL					
	190	200	210	220	230	240
m125.pep	240	250	260	270	280	299
	FTGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVT					
g125	FTGETDVAKILLGAGLGITGILAVVLSTVTTTFLDTYSAGASANNISARFAEIPVAVGVT					
	250	260	270	280	290	300
m125.pep	300	310	320	330	340	
	LIGTVLAVMLPVTEYENFLLLIGSVFAPMAGGFDCRLFRLETAX					
g125	LIRTVLAVMLPVTEYKNFLLLIRSVFGPMAGGFDCRLFCCLKTAX					
	310	320	330	340		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 461>:

a125.seq  
 1 ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCGCCGCCA TCGGGCTGAT  
 51 TTGGTTCGGC GCGGCGGTAT CGATTGCCGA AATCAGCACG GGTACACTGC  
 101 TTGCGCCTTT GGGCTGGCAG CGCGTCTGG CGGCTCTGCT TTTGGGTCAT



```

151 GCCGTCGGCG GCGCGCTGTT TTTTGGCGCG GCGTATATCG GCGCACTGAC
201 CCGACGCAGC TCGATGGAAA GCGTGC GCCT GTCGTTCCGC AAACGCGGTT
251 CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
301 GTGATGATTT ACGCCGGCGC AACGGTCAGC TCCGCTTTGG GCAAAGTGTT
351 GTGGGACGGC GAATCTTTTG TCTGGTGGGC ATTGGCAAAC GGCGCGCTGA
401 TTGTGCTGTG GCTGTTTTC GGCGCACGCA AAACAGGCGG GCTGAAAACC
451 GTTTCGATGC TGCTGATGCT GTTGGCGGTT CTGTGGCTGA GTGCCGAAGT
501 CTTTTCCACG GCAGGCAGCA CCGCCGCACA GGTTTCAGAC GGCATGAGTT
551 TCGGAACGGC AGTCGAGCTG TCCGCCGTGA TGCCGCTTTC TTGGCTGCCG
601 CTGGCCGCCG ACTACACGCG CCACGCGCGC CGCCCGTTTG CGGCAACCCCT
651 GACGGCAACG CTCGCCTACA CGCTGACCGG CTGCTGGATG TATGCCTTGG
701 GTTTGGCAGC GCGGTTGTTC ACCGAGAGAA CCGACGTGGC AAAAATCCTG
751 CTGGGCGCAG GTTTGGGTGC GGCAGGCATT TTGGCGGTCG TCCTGTGCGAC
801 CGTTACCACC ACTTTCTCG ATGCCTACTC CGCCGGCGTA AGTGCCAAACA
851 ATATTTCCGC CAAACTTTTCG GAAATACCCA TCGCCGTGCG CGTCGCCGTT
901 GTCGGCACAC TGCTTGCCGT CCTCCTGCCC GTTACCGAAT ATGAAAACCTT
951 CCTGCTGCTT ATCGGCTCGG TATTGCGGCC GATGGCG.GC GGTTTTGATT
1001 GCCGACTTTT TCGTCTTGAA ACGGCGTGA

```

This corresponds to the amino acid sequence <SEQ ID 462; ORF 125.a>:

```

a125.pep
1   MSGNASSPSS SAAIGLIWFG AAVSIAEIST GTLLAPLGWQ RGLAALLLGH
51  AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFSA NMLQLAGWTA
101 VMYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT
151 VSMLLMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP
201 LAADYTRHAR RPFATLTAT LAYTLTGCM YALGLAAALF TGETDVAKIL
251 LGAGLGAAGI LAVVLSTVTT TFLDAYSAGV SANNISAKLS EIPIAVAVAV
301 VGTLLAVLLP VTEYENFLLL IGSVFAPMAX GFDCRLFRLE TA*

```

m125/a125 95.6% identity in 342 aa overlap

```

          10      20      30      40      50      60
m125.pep  MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQ RGLAALLLGHAVGGALFFAA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a125      MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQ RGLAALLLGHAVGGALFFAA
          10      20      30      40      50      60

          70      80      90     100     110     120
m125.pep  AYIGALTGRSSMESVRLSFGKRGSVLFVANMLQLAGWTAVMIYAGATVSSALGKVLWDG
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a125      AYIGALTGRSSMESVRLSFGKRGSVLFVANMLQLAGWTAVMIYAGATVSSALGKVLWDG
          70      80      90     100     110     120

          130     140     150     160     170     180
m125.pep  ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQVSD
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a125      ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQVSD
          130     140     150     160     170     180

          190     200     210     220     230     240
m125.pep  GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCMYALGLAAALF
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a125      GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCMYALGLAAALF
          190     200     210     220     230     240

          250     260     270     280     290     300
m125.pep  TGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVTL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a125      TGETDVAKILLGAGLGAAGILAVVLSTVTTTFLDAYSAGVSANNISAKLSEIPIAVAVAV
          250     260     270     280     290     300

          310     320     330     340
m125.pep  IGTVLAVMLPVTEYENFLLLLIGSVFAPMAGGFDCRLFRLETAX
          :||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a125      VGTLLAVLLPVTEYENFLLLLIGSVFAPMAXGFDCRLFRLETAX

```

310 320 330 340

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 463>:

g126.seq

```

1   AtgccgctcTG AAaccCcaaa ggcACGCCGC CGGCTTTCAG ACGGCATCGC
51  GTCCGACAAC CATACCAAAG AATCCATCAT GCTCACCctg tacggcGAAA
101 CTTTCCCTTC GCGGCTGCTg ctcggcacgG cggcctacCC GACCCCTGAA
151 ATCCTCAAAC AATCCGTCGG AACCGCCCGG CCCGCGATGA ttaccGTCTC
201 GCTGCGCCGC ACGGGATGCG GCGGCGAGGC GCACGGTCAG GGGTTTTGGT
251 CGCTGCTTCA AGAAACCGGC GTTCCCGTCC TGCCGAACAC GGCAGGCTGC
301 CAAAGCGTGC AGGAAGCGGT AACGACGGCG CAAATGGCGC GCGAAGTGTT
351 TGAAACCGAT TGGATAAAAT TGGAATCAT CATCGCAGCAC GACACCTTGC
401 AGCCGGACGT GTTCCAACCT GTCGAAGCGG CGGAAATCCT GATTAAAGAC
451 GGCTTCAAAG TGCTGCCTTA TTGCACCGAA GACCTGATTG CCTGCCGCCG
501 CCTGCTCGAT GCGGGCTGTC AGGCGTTGAT GCCGTGGGCG GCTCCCATCG
551 GCACGGGTTT GGGGGCGGTT CACGCCTATG CGCTCAAAAT CCTGCGCGAA
601 CGCCTGCCCG ACACGCCGCT GATTATCGAC GCGGGCTTGG GTTTGCCTTC
651 CCAAGCGGCA CAAGTGATGG AATGGGGTTT TGACGGCGTA TTGTTAAACA
701 CCGCCGTTC CCGCAGCGGC GACCCCGTCA ACATGGCGCG CGCCTTCGCA
751 CTCGCCGTCG AATCCGGACG GCTGGCATT GAAGCCGGGC CGGTCGAAGC
801 GCGAACCAAA GCCCAAGCCA GCACGCCGAC AGTCGGACAA CCGTTTTGGC
851 ATTCGGCGGA ATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 464; ORF 126.ng>:

g126.pep

```

1   MPSETPKARR RLSDGIASDN HTKESIMLT YGETFPSRL LGTAAYPTPE
51  ILKQSVRTAR PAMITVSLRR TCGGGEAHGQ GFWSLLQETG VPLPNTAGC
101 QSVQEAHTTA QMAREVFETD WIKLELIGDD DTLQPDVFQL VEAAILIKD
151 GFKVLPYCTE DLIACRRLLD AGCQALMPWA APIGTGLGAV HAYALKILRE
201 RLPDTPLIID AGLGLPSQAA QVMEWFGDV LLNTAVSRSG DPVNMARAF
251 LAVESGRLEF EAGPVEARTK AQASTPTVGO PFWHSABY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 465>:

m126.seq (partial)

```

1   ..CACTATACAA AGGAACCCAT TATGCTCACC CTATACGGCG AAACTTTCCC
51  CTCGCGGCTG CTGCTCGGCA CGGCTGCCTA CCCGACCCCG GAAATCCTCA
101 AACAATCCAT CCAAACCGCC CAGCCTGCGA TGATTACCGT CTCGCTGCGC
151 CGCGCGGGAA GCGGCGGCGA GCGCACGGT CAGGGGTTTT GGTGCTGCT
201 TCAAGAAACC GCGTTCCTCG TCCTGCCGAA CACGGCAGGC TGCCAAAGCG
251 TGCAGGAAGC GGTAAACGAC GCGCAAATGG CGCGCGAAGT GTTTGAAACC
301 GATTGGATAA AATTGGAAC CTATCGGAGAT GACGACACCT TGCAGCCGGA
351 TGTGTTCCAG CTTGTGCAAG CCGCGGAAAT CCTGATTAAA GACGGCTTCA
401 AAGTGCTGCC TTATTGCACC GAAGACCTGA TTGCCTGCCG CCGCCTGCTC
451 GACGCGGGCT GTCAGGCGTT GATGCCGTGG GCGGCCCCGA TCGGCACGGG
501 TTTGGGCGCG GTTACGCGCT ACGCGTTGAA CGTCTGCGC GAACGCCTGC
551 CCGACACGCC GCTGATTATC GACGCGGGCT TGGGTTTGCC CTCACAGGCG
601 GCACAAGTGA TGGAATGGGG CTTTGACGGC GTGCTTTTGA ATACTGCCGT
651 TTCCCGCAGC GCGGATCCGG TCAATATGGC ACGCGCCTTC GCACTCGCCG
701 TCGAATCCGG ACGGCTGGCA TTTGAAGCCG GACCGGTCGA AGCACGCGAC
751 AAAGCGCAAG CCAGCACGCC GACAGTCGGA CAACCGTTTT GGCATTGCGC
801 GGAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 466; ORF 126>:

m126.pep (partial)

```

1   ..HYTKEPIMLT LYGETFPSRL LLGTAAYPTP EILKQSIQTA QPAMITVSLR
51  RAGSGGEAHG QGFWSLLQET GVPVLPNTAG QSVQEAHTT AQMAREVPET

```

```

101 DWIKLELIGD DDTLQPDVFQ LVEAAEILIK DGFKVLPYCT EDLIACRRL
151 DAGCQALMPW AAPIGTGLGA VHAYALNVL RERLPDTPLII DAGLGLPSQA
201 AQVMEWGF DG VLLNTAVSRS GDPVNMARAF ALAVESGR LA FEAGPVEARD
251 KAQASTPTVG QPFWHS AEY*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 126 shows 95.9% identity over a 269 aa overlap with a predicted ORF (ORF 126.ng) from *N. gonorrhoeae*:

m126/g126

```

                                10      20      30      40
m126.pep                      HYTK EPI MLTLYGETFPSRLL LGTAAYPTPEILKQSIQTAQ
                                ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      MPSETPKARRRLSDGIASDNHTKESIMLTLYGETFPSRLL LGTAAYPTPEILKQSVRTAR
                                10      20      30      40      50      60

                                50      60      70      80      90      100
m126.pep      PAMITVSLRRAGSGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      PAMITVSLRRTGCGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
                70      80      90      100      110      120

                                110      120      130      140      150      160
m126.pep      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                130      140      150      160      170      180

                                170      180      190      200      210      220
m126.pep      APIGTGLGAVHAYALNVLRRERLPDTPLIIDAGLGLPSQAAQVMEWGF DG VLLNTAVSRSG
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      APIGTGLGAVHAYALKILRRERLPDTPLIIDAGLGLPSQAAQVMEWGF DG VLLNTAVSRSG
                190      200      210      220      230      240

                                230      240      250      260      270
m126.pep      DPVNMARAFALAVESGR LA FEAGPVEARDKAQASTPTVGQPFWHS AEYX
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      DPVNMARAFALAVESGR LA FEAGPVEARTKAQASTPTVGQPFWHS AEYX
                250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 467>:

```

a126.seq
1   TTGTTAATCC ACTATACAAA GGAACCCATT ATGCTCACCC TGTACAGCGA
51  AACTTTCCTT TCGCGGCTGC TGCTCGGCAC AGCCGCCTAC CCGACCCCTG
101 AAATCCTCAA ACAATCCGTC CGAACC GCCCGCGAT GATTACCGTC
151 TCGCTGCGCC GCGCGGGATG CCGCGGCGAG GCGCACGGTC AGGGGTTTGT
201 GTCGCTGCTT CAAGAAACCG GCGTTCCTCGT CCTGCCGAAC ACGGCAGGCT
251 GCCAAAGCGT GCAGGAAGCG GTAACGACGG CGCAAATGGC GCGCGAAGTG
301 TTTGAAACCG ATTGGATTAA ACTCGAACTC ATCGGCGACG ACGACACCTT
351 GCAGCCGGAT GTGTTCCAAC TTGTCGAAGC GGCGGAAATC CTGATTAAAG
401 ACGGCTTCAA AGTGCTGCCT TATTGCACCG AAGACCTGAT TGCCTGCCGC
451 CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG ATGCCGTGGG CGGCCCGGAT
501 CGGCACGGGT TTGGGCGCGG TTCACGCTA CGCGTTGAAC GTCCTGCGCG
551 AACGCCTGCC CGACAGCCG CTGATTATCG ACGCGGGCTT GGGTTTGGCC
601 TCACAGCGCG CACAAGTGAT GGAATGGGGC TTTGACGCGG TGCTTTTGAA
651 TACTGCCGTT TCCGCGAGCG GCGATCCGGT CAATATGGCA CGCGCCTTCG
701 CACTCGCCGT CGAATCCGGA CGGCTGGCAT TTGAAGCCGG ACCGGTCGAA
751 GCACGCGACA AAGCGCAAGC CAGCAGCCG ACAGTCGGAC AACCGTTTGT
801 GCATTGCGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 468; ORF 126.a>:

```
a126.pep
  1  LLIHYTKEPI MLTLYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV
 51  SLRRAGCGGE AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTQAQMAREV
101  FETDWIKLEL IGDDDTLQPD VFQLVAAEI LIKDGFKVLP YCTEDLIACR
151  RLLDAGCQAL MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP
201  SQAAQVMEWG FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE
251  ARDKAQASTP TVGQPFWWSA EY*
```

m126/a126 98.1% identity in 269 aa overlap

```

              10      20      30      40      50
m126.pep      HYTKEPIMLTLYGETFP SRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGE
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           LLIHYTKEPIMLTLYSETFP SRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRAGCGGE
              10      20      30      40      50      60

              60      70      80      90      100     110
m126.pep      AHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPD
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           AHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPD
              70      80      90      100     110     120

              120     130     140     150     160     170
m126.pep      VFQLVAAEILIKDGFKVLPYCTEDLIACRRLDAGCQALMPWAAPIGTGLGAVHAYALN
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           VFQLVAAEILIKDGFKVLPYCTEDLIACRRLDAGCQALMPWAAPIGTGLGAVHAYALN
              130     140     150     160     170     180

              180     190     200     210     220     230
m126.pep      VLRERLPDTP LIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMA RAFALAVESG
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
a126           VLRERLPDTP LIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMA RAFALAVESG
              190     200     210     220     230     240

              240     250     260     270
m126.pep      RLAFEAGPVEARDKAQASTPTVGQPFWWSAEYX
              |||||:|||||:|||||:|||||:|||||:|||||
a126           RLAFEAGPVEARDKAQASTPTVGQPFWWSAEYX
              250     260     270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 469>:

```
g126-1.seq
  1  ATGCTCACCC TGTACGGCGA AACTTTCCCT TCGCGGCTGC TGCTCGGCAC
 51  GGCCCGCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACC GCCC
101  GGCCCGCGAT GATTACCGTC TCGCTCGGCC GCACGGGATG CGGCGGCGAG
151  GCGCAGCGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201  CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251  CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAACTC
301  ATCGCGCAGC ACGACACCTT GCAGCCGGAC GTGTTCCAAC TCGTCGAAGC
351  GGCGGAAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
401  AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ATGCGGGCTG TCAGGCGTTG
451  ATGCGTGGG CCGCTCCCAT CCGCAGGGT TTGGGGGCGG TTCACGCCTA
501  TGCGCTCAA ATCCTGCGCG AACGCTGCC CGACACGCCG CTGATTATCG
551  ACGCGGGCTT GGGTTTGCTT TCCCAAGCGG CACAAGTGAT GGAATGGGGT
601  TTTGACGGCG TATTGTAAAC CACCGCCGTT TCCCGCAGCG GCGACCCCGT
651  CAACATGGCG CGCGCTTCG CACTCGCCGT CGAATCCGCG CGGCTGGCAT
701  TTGAAGCCGG GCCGCTCGAA GCGCGAACCA AAGCCCAAGC CAGCACGCCG
751  ACAGTCGGAC AACC GTTTTG GCATTGCGCG GAATATTGA
```

This corresponds to the amino acid sequence <SEQ ID 470; ORF 126-1.ng>:

```
g126-1.pep
  1  MLTLYGETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRTGCGGE
```

```

51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQVLEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALK ILRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARTKAQASTP
251 TVGQPFWWSA EY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 471>:

m126-1.seq

```

1 ATGCTCACCC TATACGGCGA AACTTTCCCC TCGCGGCTGC TGCTCGGCAC
51 GGCTGCCTAC CCGACCCCGG AAATCCTCAA ACAATCCATC CAAACCGCCC
101 AGCCTGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGAAG CGGCGGCGAG
151 GCGCAGGCTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251 CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAAGTC
301 ATCGGAGATG ACGACACCTT GCAGCCGGAT GTGTTCAGC TTGTGGAAGC
351 GCGCGAAATC CTGATTAAAG ACGGCTTCAA AGTGTGCTT TATTGCACCG
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451 ATGCCGTGGG CGGCCCGGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501 CGCGTTGAAC GTCCTGCGCG AACGCCTGCC CGACACGCCG CTGATTATCG
551 ACGCGGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601 TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
651 CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701 TTGAAGCCGG ACCGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
751 ACAGTCGGAC AACCCTTTTG GCATTCGGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 5; ORF 126-1>:

m126-1.pep

```

1 MLTYGETFP SRLLLGTAAY PTPEILKQSI QTAQPAMITV SLRRAGSGGE
51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQVLEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWWSA EY*

```

m126-1/g126-1 96.9% identity in 262 aa overlap

	10	20	30	40	50	60
m126-1.pep	MLTYGETFP SRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGEAHGQGFWSLL					
g126-1	MLTYGETFP SRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRTGCGGEAHGQGFWSLL					
	10	20	30	40	50	60
m126-1.pep	QETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPDVFQVLEAAEI					
g126-1	QETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPDVFQVLEAAEI					
	70	80	90	100	110	120
m126-1.pep	QETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPDVFQVLEAAEI					
g126-1	QETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPDVFQVLEAAEI					
	70	80	90	100	110	120
m126-1.pep	LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALNVLRERLPDTP					
g126-1	LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALKILRERLPDTP					
	130	140	150	160	170	180
m126-1.pep	LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALNVLRERLPDTP					
g126-1	LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALKILRERLPDTP					
	130	140	150	160	170	180
m126-1.pep	LIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMA RAFALAVESG RLAFEAGPVE					
g126-1	LIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMA RAFALAVESG RLAFEAGPVE					
	190	200	210	220	230	240
m126-1.pep	LIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMA RAFALAVESG RLAFEAGPVE					
g126-1	LIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMA RAFALAVESG RLAFEAGPVE					
	190	200	210	220	230	240
m126-1.pep	ARDKAQASTPTVGQPFWWSAEYX					
g126-1	ARTKAQASTPTVGQPFWWSAEYX					
	250	260				
m126-1.pep	ARDKAQASTPTVGQPFWWSAEYX					
g126-1	ARTKAQASTPTVGQPFWWSAEYX					
	250	260				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 473>:

a126-1.seq

```

1 ATGCTCACCC TGTACAGCGA AACTTTCCCT TCGCGGCTGC TGCTCGGCAC
51 AGCCGCTTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACCGCCC
101 GGCCGCGCAT GATTACCGTC TCGCTGCGCC GCGCGGGATG CGGCGGCGAG

```

```
151 GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251 CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATTAA ACTCGAACTC
301 ATCGGCGACG ACGACACCTT GCAGCCGGAT GTGTTCCAAC TTGTCGAAGC
351 GCGCGAAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451 ATGCCGTGGG CGGCCCGGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501 CGCGTTGAAC GTCCTGCGCG AACGCCTGCC CGACACGCCG CTGATTATCG
551 ACGCGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601 TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
651 CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701 TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
751 ACAGTCGGAC AACCGTTTTG GCATTGCGCG GAATATTGA
```

This corresponds to the amino acid sequence <SEQ ID 474; ORF 126-1.a>:

a126-1.pep

```
1 MLTLYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRAGCGGE
51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDNIKLEL
101 IGDDDTLQPD VFQVLEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMWEG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWWSA EY*
```

a126-1/m126-1 98.1% identity in 262 aa overlap

a126-1.pep	10	20	30	40	50	60
	MLTLYSETFP	SRLLLGTAAY	PTPEILKQSV	RTARPAMITV	SLRRAGCGGE	AHGQGFWSLL
m126-1	10	20	30	40	50	60
	MLTLYGETFP	SRLLLGTAAY	PTPEILKQSI	OTAQPMITV	SLRRAGSGGE	AHGQGFWSLL
a126-1.pep	70	80	90	100	110	120
	QETGVPVLP	NTAGCQSVQEA	VTTAQMAREV	FETDNIKLEL	IGDDDTLQPD	VFQVLEAAEI
m126-1	70	80	90	100	110	120
	QETGVPVLP	NTAGCQSVQEA	VTTAQMAREV	FETDNIKLEL	IGDDDTLQPD	VFQVLEAAEI
a126-1.pep	130	140	150	160	170	180
	LIKDGFKVLP	YCTEDLIACR	RLLDAGCQAL	MPWAAPIGT	GLGAVHAYAL	NVLRERLPDTP
m126-1	130	140	150	160	170	180
	LIKDGFKVLP	YCTEDLIACR	RLLDAGCQAL	MPWAAPIGT	GLGAVHAYAL	NVLRERLPDTP
a126-1.pep	190	200	210	220	230	240
	LIIDAGLGLP	SQAAQVMWEG	FDGVLLNTAV	SRSGDPVNMA	RAFALAVESG	RLAFEAGPVE
m126-1	190	200	210	220	230	240
	LIIDAGLGLP	SQAAQVMWEG	FDGVLLNTAV	SRSGDPVNMA	RAFALAVESG	RLAFEAGPVE
a126-1.pep	250	260				
	ARDKAQASTP	TVGQPFWWSA	EYX			
m126-1	250	260				
	ARDKAQASTP	TVGQPFWWSA	EYX			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 475>:

g127.seq

```
1 ATGGAATAT GGAATATGTT GAACACTTGG CCCGATGCCG TCCCAGATACG
51 CGCGGAGGCG GCCGAATCCG TGGCGGCGGT CGCGGCTTTG CTGCTGGCGC
101 GCGCCCTTCT GTTGAATATC CACTTCAGAC GGCATCCGGA TTTCGGCATC
151 GAAAGCAAGC GCGGTTTTTT GGTGGCCAGC CGCAATATAA CGCTGCTTTT
201 GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GCGCGAAATT CAAACGCTGG
251 CTTTGTCGAT GTTTCGGGTG GCGGCGCGCG TCGTCGTGGC GACAAAAGAA
301 CTGATTATGT GTCTGTCGGG CAGTATTTTA aggtctGCCA CCCAGCAATA
351 CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
401 ACATCAATCT GTTGAACACG CTGATGATGC AGGTGGGTCC GAACCCCTTG
451 GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCAACA GCCTGTTGTT
501 GAGCCACCCC GTGCGCCGCG ACAATATTTT GGGCGACTAT GTCATCCATA
```

```

551 CGGTCGAAAT CCCCGTTCCC ATCCATTGG ATTCGGATGA AGCCGTATGC
601 CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
651 TCAGCGGTAT TTGGAAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCCG
701 CCGCCAGGCC GCGCGTTACC CGCGTACCGT ACGACGACAA GGCATACCGC
751 ATCATCGTCC GCTTCGCCTC CCCCGTTTCA AAGCGGCTGG AAATCCAACA
801 GGCGGTTATG GACGAATTTT TGCGCGTACA ATACCGCCTG TTAAATCATC
851 CCGCCGgctc cgAAACACTT TAA

```

This corresponds to the amino acid sequence <SEQ ID 476; ORF 127.ng>:

```

g127.pep
  1 MEIWNMLNTW PDAVPIRAEA AESVAAVAAL LLARALLLNI HFRRHPDFGI
 51 ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAAVVVATKE
101 LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGPNNPL
151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201 RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251 IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 477>:

```

m127.seq
  1 ATGGAAATAT GGAATATGTT GGACACTTGG CTCGGTGCCG TCCCGATACG
 51 TCGCGAGGCG GTCGAATCCG TGGCGGCGGT TGCGGCTTTG CTGCTGGCGC
101 GCGCCCTTCT GTTGAATATC CACTTCAAAC GGCATCCGGA TTTCGGCATC
151 GAAAGCAAGC GGCGGTTTTT GGTGCCAGC CGCAATATAA CGTGCTTTTT
201 GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GGCGCAAATC CAAACGCTGG
251 CTTTGTGCGT GTTTGCGGTG GCGGCGGCGG TCGTCGTGGC GACGAAGGAA
301 CTGATTATGT GTCTGTCGGG CAGTATTTTA AGGTCTGCCA CCCAGCAATA
351 CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
401 ACATCAACCT GTTGAACACG CTGATGATGC AGGTCGGTCC GAACCCCTTG
451 GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCCAACA GCCTGTTGTT
501 GAGCCACCCC GTGCGCCGCG ACAATATTTT GGGCGACTAT GTCATCCATA
551 CGGTCGAAAT CCCCGTTCCC ATCCATTGG ATTCGGATGA AGCCGTATGC
601 CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
651 CCAACGGsAT TTGGAAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCCG
701 CCGCCAGACC GCGCGTTACC CGCGTGCCGT ACGATGACAA GGCATACCGC
751 ATCATCGTCC GCTTCGCTTC CCCCGTTTCA AAGCGGCTGG AAATCCAACA
801 GGCGGTTATG GACGAATTTT TGCGCGTACA ATACCGCCTG TTAAATCACC
851 CCGCCGCTC CGAAACACTT TAA

```

This corresponds to the amino acid sequence <SEQ ID 478; ORF 127>:

```

m127.pep
  1 MEIWNMLDTW LGAVPIRAEA VESVAAVAAL LLARALLLNI HFKRHPDFGI
 51 ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAAVVVATKE
101 LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGPNNPL
151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201 RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251 IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 127 shows 97.9% identity over a 290 aa overlap with a predicted ORF (ORF 127.ng) from *N. gonorrhoeae*:

```

m127/g127
      10      20      30      40      50      60
m127.pep MEIWNMLDTWLGAVPIRAEAVESVAAVAALLLARALLLNIHFKRHPDFGIESKRRFLVAS
          |||||:|  |||||:|  |||||:|  |||||:|  |||||:|  |||||:|
g127      MEIWNMLNTWPDVPIRAEAESVAAVAALLLARALLLNIHFRRHPDFGIESKRRFLVAS
          10      20      30      40      50      60

      70      80      90     100     110     120
m127.pep RNITLLLVLFSLAFIWSAQIQTALSMFAVAAAVVVATKELIMCLSGSILRSATQQYSVG
          |||||:|  |||||:|  |||||:|  |||||:|  |||||:|  |||||:|
g127      RNITLLLVLFSLAFIWSAQIQTALSMFAVAAAVVVATKELIMCLSGSILRSATQQYSVG

```

	70	80	90	100	110	120
	130	140	150	160	170	180
m127.pep	DYIEINGLRGRVVDINLLNTLMMQVGNPLVQLAGTTVSFPNSLLLSHPVRRDNILGDY					
g127	DYIEINGLRGRVVDINLLNTLMMQVGNPLVQLAGTTVSFPNSLLLSHPVRRDNILGDY					
	130	140	150	160	170	180
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLDSDCAVCRLLKAVLEPLCAPYIPAIQRXLENVQAEKLFITPAARPRVT					
g127	VIHTVEIPVPIHLDSDCAVCRLLKAVLEPLCAPYIPAIQRYLENVQAEKLFITPAARPRVT					
	190	200	210	220	230	240
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLLNHPAGSETLX					
g127	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLLNHPAGSETLX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 479>:

a127.seq	1	ATGGAAATAT	GGAATATGTT	GGACACTTGG	CTCGGTGCCG	TCCCGATACG
	51	TGCGGAGGCG	GTCGAATCCG	TGCGGGTGGT	CGCGGCTTTG	CTGCTGGCGC
	101	GCGCCCTTCT	GTTGAATATC	CACTTCAAAC	GGCATCCGGA	TTTCGGCATC
	151	GAAAGCAAGC	GGCGGTTTTT	GGTTGCCAGC	CGCAATATAA	CGCTGCTTTT
	201	GGTGCTGTTT	TCGCTGGCAT	TTATCTGGTC	GGCGCAAATC	CAAACGCTGG
	251	CTTTGTTCGAT	GTTTGCAGTG	GCGGCGGCGG	TCGTCGTGGC	GACGAAGGAA
	301	CTGATTATGT	GTCTGTGCGG	CAGCATTTTA	AGGTCTGCCA	CCCAGCAATA
	351	CTCGGTCGGC	GACTATATCG	AAATCAACGG	CCTGCGCGGG	CGCGTGGTCC
	401	ACATCAACCT	GTTGAACACG	CTGATGATGC	AGGTCGGTCC	GAACCCCTTG
	451	GTCGGACAGC	TTGCGGGAAC	CACCGTTTCT	TTCCCCAACA	GCCTGTTGTT
	501	GAGCCACCCC	GTGCGCCGCG	ACAATATTTT	GGGCGACTAC	GTCATCCATA
	551	CGGTCGAAAT	CCCGGTTCCC	ATCCATTTGG	ATTCGGATGA	AGCCGTATGC
	601	GCTCTGAAAG	CCGTAATCGA	GCCCTTGTCG	GCGCCCTACA	TCCCGGCCAT
	651	CCAACGGCAT	TTGAAAACG	TGCAGGCGGA	AAAACGTGTT	ATCACGCCCG
	701	CCGCCAAACC	GCGCGTTACC	CGCGTGCCGT	ACGATGACAA	GGCATACCGC
	751	ATCATCGTCC	GCTTCGCCTC	CCCCGTTTCA	AAGCGGCTGG	AAATCCAACA
	801	GGCGGTTATG	GACGAATTTT	TGCGCGTACA	ATACCGCCTG	TTAAATTACC
	851	CCGCCGGCTC	CGAAACACTT	TAA		

This corresponds to the amino acid sequence <SEQ ID 480; ORF 127.a>:

a127.pep	1	MEIWNMLDTW	LGAVPIRAEA	VESVAVVAAL	LLARALLLNI	HFKRHPDFGI
	51	ESKRRLVAS	RNITLLLVLF	SLAFIWSAQI	QTLALSMFAV	AAAVVVATKE
	101	LIMCLSGSIL	RSATQQYSVG	DYIEINGLRG	RVVDINLLNT	LMMQVGNPL
	151	VGQLAGTTVS	FPNSLLLSHP	VRRDNILGDY	VIHTVEIPVP	IHLDSDEAVC
	201	RLKAVLEPLC	APYIPAIQRH	LENVQAEKLF	ITPAKPRVT	RVPYDDKAYR
	251	IIVRFASPV	KRLEIQQAVM	DEFLRVQYRL	LNYPAGSETL	*

m127/a127 98.6% identity in 290 aa overlap

	10	20	30	40	50	60
m127.pep	MEIWNMLDTWLGAVPIRAEA	VESVAVVAALLARALLLNI	HFKRHPDFGIESKRRLVAS			
a127	MEIWNMLDTWLGAVPIRAEA	VESVAVVAALLARALLLNI	HFKRHPDFGIESKRRLVAS			
	10	20	30	40	50	60
	70	80	90	100	110	120
m127.pep	RNITLLLVLFSLAFIWSAQI	QTLALSMFAVAAAVVATKEL	IMCLSGSILRSATQQYSVG			
a127	RNITLLLVLFSLAFIWSAQI	QTLALSMFAVAAAVVATKEL	IMCLSGSILRSATQQYSVG			
	70	80	90	100	110	120



	130	140	150	160	170	180
m127.pep	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY					
a127	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY					
	130	140	150	160	170	180
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRXLENVQAEKLFITPAARPRVT					
a127	VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRHLENVQAEKLFITPAARPRVT					
	190	200	210	220	230	240
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLNHPAGSETLX					
a127	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLNYPAGSETLX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 481>:

g128.seq

```

1  atgattgaca  acgCActgct  ccacttgggc  gaagaaccCC  GTTTTaatca
51  aatccaaacc  gaagACAtca  AACCCGCCGT  CCAAACCGCC  ATCGCCGAAG
101 CGCGCGGACA  AATCGCCGCC  GTCAAAGCGC  AAACGCACAC  CGGCTGGGCG
151 AACACCGTCG  AGCGTCTGAC  CGGCATCACC  GAACGCGTCG  GCAGGATTGT
201 GGGCGTCGTG  TCCCATCTCA  ACTCCGTCGT  CGACACGCCC  GAACTGCGCG
251 CCGTCTATAA  CGAACTGATG  CCTGAAATCA  CCGTCTTCTT  CACCGAAATC
301 GGACAAGACA  TCGAACTGTA  CAACCGCTTC  AAAACCATCA  AAAATTCCCC
351 CGAATTGCA  ACGCTTTCCT  CCGCACAAAA  AACCAAGCTC  GATCACGACC
401 TGCGCGATTT  CGTATTGAGC  GGCGCGGAAC  TGCCGCCCGA  ACGGCAGGCA
451 GAACTGGCAA  AACTGCAAAC  CGAAGGCGCG  CAACTTTCCT  CCAAATTCTC
501 CCAAAACGTC  CTAGACGCGA  CCGACGCGTT  CGGCATTAC  TTTGACGATG
551 CCGCACCGCT  TGCCGGCATT  CCCGAAGACG  CGCTCGCCAT  GTTTGCCGCC
601 GCCGCGCAAA  GCGAAGGCAA  AACAGGTTAC  AAAATCGGCT  TGCAGATTCC
651 GCACTACCTT  GCCGTTATCC  AATACGCCGG  CAACCGCGAA  CTGCGCGAAC
701 AAATCTACCG  CGCCTACGTT  ACCCGTGCCA  GCGAACTTTC  AAACGACGGC
751 AAATTCGACA  ACACCGCCAA  CATCGACCGC  ACGCTCGAAA  ACGCATTGAA
801 AACCGccaaa  cTGCTCGGCT  TTAAAAATTA  CGCCGAATTG  TCGCTGGCAA
851 CCAAAATGGC  GGACACGCCC  GAACAGGTTT  TAAACTTCCT  GCACGACCTC
901 GCCCGCCGCG  CCAAACCCTA  CGCCGAAAAA  GACCTCGCCG  AAGTCAAAGC
951 CTTCGCCCCG  GAACACCTCG  GTCTCGCCGA  CCCGCAGCCG  TGGGACTTGA
1001 GCTACGCCGG  CGAAAACTG  CGCGAAGCCA  AATACGCATT  CAGCGAAACC
1051 GAAGTCAAAA  AATACTTCCC  CGTCGGCAAA  GTTCTGGCAG  GCCTGTTTCG
1101 CCAAATCAAA  AAATCTACG  GCATCGGATT  CGCCGAAAAA  ACCGTTCCCC
1151 TCTGGCACA  AGACGTGCGC  TATTTTGAAT  TGCAACAAAA  CGGCAAAACC
1201 ATCGGCGGCG  TTTATATGGA  TTTGTACGCA  CGCGAAGGCA  AACGCGCGCG
1251 CGCGTGATG  AACGACTaca  AAGGCCGCCG  CCGCTTTGCC  GACGgcacGC
1301 TGCAACTGCC  CACCGCCTAC  CTCGTCTGCA  ACTTCGCCCC  GCCCGTCGGC
1351 GGCAAAGAAG  CGCGTTTAAG  CCACGACGAA  ATCCTCACC  TCTTCCACGA
1401 AacCGGCCAC  GGACTGCACC  ACCTGCTTAC  CCAAGTGGAC  GAACTGGGCG
1451 TGTCGGGCAT  CAacggcgtA  GAATGGGACG  CGGTGCAACT  GCCAGCCAG
1501 TTTATGGA  ACTTCGTTTG  GGAATACAAT  GTATTGGCAC  AAATGTCCGC
1551 CCACGAAGAA  AccgGCGAGC  CCCTGCCGAA  AGAACTCTTC  GACAAAAATG
1601 TcgcCGCCAA  AAATTTCCAG  CGCGGTATGT  TCCTCGTCCG  GCAAATGGAG
1651 TTCGCCCTCT  TCGATATGAT  GATTTACAGT  GAAAGCGACG  AATGCCGTCT
1701 GAAAAACTGG  CAGCAGGTTT  TAGACAGCGT  GCGCAAAGAA  GTcGCCGTCA
1751 TCCAACCGCC  CGAATACAAC  CGCTTCGCCA  ACAGCTTCGG  CCacatctTC
1801 GCcggcGGCT  ATTCCGCGAG  CTATTACAGC  TACGCATGGG  CCGAAGTCct
1851 cAGCACCGAT  GCCTACGCCG  CCTTTGAAGA  AAGcGACGac  gtcGCCGCCA
1901 CAGGCAAACG  CTTCTGGCAA  GAAAtccttg  ccgtcggcgg  ctCCCGCAGC
1951 gcgGCGGAAT  CTTCAAAGC  CTTCCGCGGA  CGCGAACCAG  GCATAGACGC

```

2001 ACTGCTGCGC CAAagcggtT TCGACAACGC gGcttgA

This corresponds to the amino acid sequence <SEQ ID 482; ORF 128.ng>:

g128.pep

```

1  MIDNALLHLG EEPFRNQIQT EDIKPAVQTA IAEARGQIAA VKAQHTGTGWA
51  NTVRLTGIT ERVGRWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 QDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDSYAGEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQONGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS ESDECRLKNW QQVLDSVRKE VAVIQPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSTD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR QSGFDNAA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 483>:

m128.seq (partial)

```

1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACC CGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCTGAC CGGCATCACC GAACCGCTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTGCGTCGC CGACACGCCG GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTTCGAC ACCCTCTCCC CCGCACAAA AACC AAACCTC AACCAC
1  TACGCCAGCG AAAA ACTGCG CGAAGCCAAA TACGCGTTCA GCGAAACCGA
51  WGTCAAAAAA TAYTTCCCYG TCGGCAAWGT ATTAAACGGA CTGTTCCGCC
101 AAmTCAAAAA ACTmTACGGC ATCGGATTTA CCGAAAAAAC yGTCCCCGTC
151 TGGCACAAG ACGTGCCTA TTKTGAATTG CAACAAAACG GCGAaMCCAT
201 AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
251 CTGGGATGAA CGACTACAAA GGCCGCCGCC GTTTTTCAGA CGGCACGCTG
301 CAAYTGCCCA CCGCCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGGCGG
351 CAGGGAAGCC CGCyTGAGCC ACGACGAAAT CCTCATCCTC TTCCACGAAA
401 CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
451 TCCGGCATCA ACGGCGTAKA ATGGGACGCG GTCGAACTGC CCAGCCAGTT
501 TATGGAAAAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
551 ACGAAGAAAC CGGcgTTCCC YTGCCGAAAG AACTCTTsGA CAAAwTGCTC
601 GCCGCCAAA ACTTCCAAsG CGGCATGTTC yTsGTCCGCG AAwTGGAGTT
651 CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
701 AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTATC
751 CAGCCGCCCG AATACAACCG CTTCCGCTTG AGCTTCGGCC ACATCTTCGC
801 AGGCGGCTAT TCCGCAGCTn ATTACAGCTA CGCGTGGGCG GAAGTATTGA
851 GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
901 GGCAAACGCT TTTGGCAGGA AATCCTCGCC GTCGGGGnAT CGCGCAGCGG
951 nGCAGAAATCC TTCAAAGCCT TCCGCGGCCG CGAACCAGAG ATAGACGCAC
1001 TCTTGCGCCA CAGCGGTTTC GACAACGCGG TCTGA

```

This corresponds to the amino acid sequence <SEQ ID 484; ORF 128>:

m128.pep (partial)

```

1  MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGTGWA
51  NTVPLTGIT ERVGRWGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
101 QDIELYNRF KTIKNSPEFD TLSPAQKTKL NH

```

//

```

1  YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGFTEKTVPV
51  WHKDVRYXEL QONGEXIGGV YMDLYAREGK RGGAWMNDYK GRRRFSDGTL
101 QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
151 SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
201 AAKNFQXGMF XVROXEFALF DMMIYSEDDE GRLKNWQOVL DSVRKKVAVI
251 QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT

```

301 GKRFWQEILA VGXSRGAES FKAFRGREPS IDALLRHSGF DNAV\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from *N. gonorrhoeae*:

m128/g128

	10	20	30	40	50	60		
g128.pep	MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQTHGTGWANTVERLTGIT							
m128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHGTGWANTVEPLTGIT							
	10	20	30	40	50	60		
g128.pep	70	80	90	100	110	120		
	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA							
m128	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD							
	70	80	90	100	110	120		
g128.pep	130	140	150	160	170	180		
	TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY							
m128	TLSPAQKTKLNH							
	130							
	//							
g128.pep						340	350	360
						YAGEKLREAKYAFSETEVKKYFPVGVKVLG		
m128						YASEKLREAKYAFSETXVKKYFPVGVXLNG		
						10	20	30
g128.pep	370	380	390	400	410	420		
	LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK							
m128	LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWMNDYK							
	40	50	60	70	80	90		
g128.pep	430	440	450	460	470	480		
	GRRRFADGTLQLPTAYLVCNFAPPVVGKEARLSHDEILTLFHETGHGLHLLTQVDELGV							
m128	GRRRFSFGTLQLPTAYLVCNFAPPVVGREARLSHDEILILFHETGHGLHLLTQVDELGV							
	100	110	120	130	140	150		
g128.pep	490	500	510	520	530	540		
	SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQGMF							
m128	SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQXGMF							
	160	170	180	190	200	210		
g128.pep	550	560	570	580	590	600		
	LVRQMEFALFDMMIYSESDCRLKNWQQVLDVSRKEVAVIQPEYNRFANSFGHIFAGGY							
m128	XVRQXEFALFDMMIYSEDDGRLKNWQQVLDVSRKKVAVIQPEYNRFALSFGHIFAGGY							
	220	230	240	250	260	270		
g128.pep	610	620	630	640	650	660		
	SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS							
m128	SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRGAESFKAFRGREPS							
	280	290	300	310	320	330		

```

          670      679
g128.pep  IDALLRQSGFDNAAX
          |||||:|||||:
m128      IDALLRHSGFDNAVX
          340

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 485>:

```

a128.seq
1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
251 CGGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAAC TC AACCACGATC
401 TCGCGGATTT CGTCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCCTG TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCAA CTGCTCGGCT TCAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTATTAAACG GACTGTTCCG
1101 CCAAATCAAA AAACCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCAACCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCTGCAAC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAACCTTCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 486; ORF 128.a>:

```

a128.pep
1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWA
51  NTVEPLTGIT ERVGRWGVV SHLNSVTDTP ELRAAYNELM PEITVFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIOYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTQLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME

```

551 FALFDMMIYS EDDEGR LKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF  
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS  
651 AAESFKAFRG REPSIDALLR HSGFDNAA\*

## m128/a128 66.0% identity in 677 aa overlap

m128.pep	10	20	30	40	50	60
	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT					
a128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT					
	10	20	30	40	50	60
m128.pep	70	80	90	100	110	120
	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
a128	ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
m128.pep	130					
	TLSPAQKTKLNH-----					
a128	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
m128.pep	-----					
a128	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV					
	190	200	210	220	230	240
m128.pep	-----					
a128	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
	250	260	270	280	290	300
m128.pep	-----					
	140	150				
	-----YASEKLREAKYAFSETXVKKYFPVGX					
a128	ARRAKPYAEKDIAEVKAFARESGLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGK					
	310	320	330	340	350	360
m128.pep	160	170	180	190	200	210
	VLNGLFAQXKKLYGIGFTEKTPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM					
a128	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
m128.pep	220	230	240	250	260	270
	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD					
a128	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
	430	440	450	460	470	480
m128.pep	280	290	300	310	320	330
	ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ					
a128	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540
m128.pep	340	350	360	370	380	390
	XGMFXVRQXEFALFDMMIYSEDEGR LKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF					
a128	RGMFLVRQMEFALFDMMIYSEDEGR LKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF					
	550	560	570	580	590	600

```

400      410      420      430      440      450
m128.pep  AGGYSAAXSYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRG
          |||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a128      AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
          610      620      630      640      650      660

460      470
m128.pep  REPSIDALLRHSGFDNAVX
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a128      REPSIDALLRHSGFDNAAX
          670

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 487>:

g128-1.seq (partial)

1	ATGATTGACA	ACGCACTGCT	CCACTTGGGC	GAAGAACCCC	GTTTAAATCA
51	AATCAAAACC	GAAGACATCA	AACCGCGCGT	CCAACCGGCC	ATCGCCGAAG
101	CGC CGCGGCA	AATCGCGGCC	GTCAAAGCGC	AAACGCACAC	CGGCTGGGCG
151	ACACCGCTCG	AGCGTCTGAC	CGGCATCACC	GAACGCGTCG	CGAGGATTTG
201	GGGCGTCTGT	TCCCATCTCA	ACTCCGTCGT	CGACACGCCC	GAACTCTGCG
251	CCGTCTATAA	CGAAGCTGATG	CCTGAAATCA	CCGTCTTCTT	CACCGAAATC
301	GGACAAGACA	TGGAAGTCTA	CAACCGCTTC	AAAAACATCA	AAAATTTCCC
351	CGAATTTGCA	ACGCTTTCCC	CCGCACAAA	AACCAAGCTC	GATCACGACC
401	TGCGCGATTT	CGTATTGAGC	GGCGCGGAAC	TGCCGCCCGA	ACCGCAGGCA
451	GAAGTGGCAA	AACTGCACAA	CGAAGGCGCG	CAACTTTCGG	CCAAATTTCTC
501	CCAAAACGCT	CTAGACGCGA	CCGACGCGTT	CGGCATTTCG	TTTGCCGATG
551	CCGCACCGCT	TGCCGGCATT	CCCGAAGACG	CGCTGCCCAT	TTTGCCGCGC
601	GCCGCGCAAA	GCGAAGGCAA	AACAGGTTAC	AAAATCGGCT	TGCAGATTCC
651	GCACTACCTT	CGCGTTATCC	AATAAGCCGG	CAACCGCGAA	CTGCGCGAAC
701	AAATCTACCG	GCCCTACGTT	ACCCGTGCCA	GCGAACTTTC	AAACGACGGC
751	AAATTCGACA	ACACCGCCAA	CATCGACCGC	ACGCTCGAAA	ACGCATTGAA
801	AACCGCCAAA	TGCTCTCGCT	TTAAAAATTA	CGCGGAATTG	TCGCTGGCAA
851	CCAAAATGGC	GGACACGCCC	GAACAGGTTT	TAAACTTCCT	GCACACCTCT
901	GCCGCGCGCG	CCAAAACCTA	GCGCGAAAAA	GACTCGCGCG	AAGTCAAAGC
951	CTTCGCGCGC	GAACACCTCG	GTCTCGCCGA	CCCGCAGCGC	TGGGACTTGA
1001	GTAACGCGCG	CGAAAAACCTG	CGCGAAGCCA	AAATACGATT	CAGCGAAACC
1051	GAAGTCAAAA	ATAACTTCCC	CGTCGGGCAA	GTCTTGGCAG	CGCTGTTCCG
1101	CCAAATCAAA	AAACTCTACG	GCATCGGATT	CGCGCAAAAA	AGCCTTCCCG
1151	TCTGGCAGCA	AGACGTGCGC	TATTTTGAAT	TGCAACAAAA	CGGCAAAACC
1201	ATCGCGCGCG	TTTATATGGA	TTGTGACGCA	CGCGAAGGCA	AACGCGGCGG
1251	CGCGTGGATG	AACGACTACA	ATGGCCGCGC	CCGCTTTGGC	GACGGCACGC
1301	TGCAACTGCC	CACCGCCTAC	CTCGTCTGCA	ACTTCGCCCC	GCCGCTCGCG
1351	GGCAAAAGAG	CGCGTTTAAAG	CCACGACGAA	ATCCTCACC	TCTTCCACGA
1401	AACCGGCCAC	GGACTGCACC	ACCTGCTTAC	CCAAGTGGAC	GAAGTGGGCG
1451	TGTCGGCGAT	CAACGCGCTA	AAA		

This corresponds to the amino acid sequence <SEQ ID 488; ORF 128-1.ng>:  
g128-1.pep (partial)

g128-1.pep (partial)

1	MIDNALLHIG	EPRFNQIKT	EDIKPAVQTA	IAEARGQIAA	VKAQHTGTWA
51	NTVERLTGTT	ERVGRIWGVV	SHLSNVVDTP	ELRAVYNELM	PEITVFFTEI
101	GQDIELYNRF	QTKSNPEFA	TLSPAQKTKL	DHDLRDFVLS	GAEPPERQA
151	ELAKLQTEGA	KLISAKFSQV	LDATDAFGIY	FDDAAPLAGI	PEDALAMFAA
201	AAQSEGKTGY	KIGLQIPHYL	AVIQYAGNRE	LREQIYRAVY	TRASELSNDG
251	KFDNTANIDR	TLENALKTAK	LLGFKNYAEL	SLATKMADTP	EQVLNFIHLH
301	ARRAKPYAEK	DLAEVKAFAR	EHLGLADPPQ	WDLSYAGEKL	REAKYAFSET
351	EVKKYFFPVG	VLAGLFAQIK	KLYGIGFAEK	TPVWVKHDVR	YFELQQNGKT
401	IGGVYMDLYA	REGKRGAWM	NLDYKGRRFA	DGTLQLPTAY	LVNCFAPPVG
451	GKEARLSHDE	ILTLFHETGH	GLHHLLTQVD	ELGVSGINGV	K

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 489>:

m128-1.seq

1 ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA  
51 AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACGCC ATCGCCGAAG  
101 CGCGCGAACA AATCGCGGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA  
151 AACACTGTCTG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG

```

201 GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCG GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACACGATC
401 TGGCGGATTT CGTCTCAGC GCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAACTGGCAA AACTGCAAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
651 AACTACCTC GCCGTATCC AATACGCCGA CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCTACGTT ACCCGCGCCA GCGAACTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGCG GGACACGCCG GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAAACCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCCG GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTTCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCAGAA AGACGTGCGC TATTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CCGGTGGATG AACGACTACA AAGGCCGCGC CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGGAA CCGCCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
1401 AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CGGTGGAATC GCCAGCCAG
1501 TTTATGGAAT ATTTCGTTT GGAATACAAT GTCTTGGCAC AAATGTCAGC
1551 CAAGSAGGAA ACCGGCGTTC CCCTGCCGAA AGAATCTTTC GACAAAATGC
1601 TCGCCGCCAA AAATTTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
1751 TCAGCCGCGC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAAAC CTTTGGGAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTCTCAAAGC CTTCCGCGGC CGCGAACCGA GCATAGACGC
2001 ACTCTTGGCG CACAGCGGTT TCGACAACGC GGTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 490; ORF 128-1>:

m128-1.pep.

```

1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGTWA
51 NTVEPLTGIT ERVGRWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSESTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLANALQTAK LLGFKNYAEL SLATKMDATP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLNLADLP WDLGYASEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTQLQPTAY LVCNFAPPVG
451 GREARLSHDE ILILFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLEKNW QQVLDVSRKK VAVIQPPEYN RFALSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESPKAFRG REPSIDALLR HSGFDNAV*

```

m128-1/g128-1 94.5% identity in 491 aa overlap

```

          10      20      30      40      50      60
g128-1.pep MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQHTGWANTVERLTGIT
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
          10      20      30      40      50      60

          70      80      90     100     110     120
g128-1.pep ERVGRWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128-1      ERVGRWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90     100     110     120

```

g128-1.pep	130	140	150	160	170	180
	TLSPAQKTKLDHDLRDFVLSGAELPPERQAEAKLQTEGAQLSAKFSQNVLDATDAFGIY					
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
g128-1.pep	190	200	210	220	230	240
	FDDAAPLAGIPEDALAMFAAAQSEGKTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV					
m128-1	FDDAAPLAGIPEDALAMFAAAQSESKTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV					
	190	200	210	220	230	240
g128-1.pep	250	260	270	280	290	300
	TRASELSNDGKFDNTANIDRTLENALKTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
	250	260	270	280	290	300
g128-1.pep	310	320	330	340	350	360
	ARRAKPYAEKDIAEVKAFAREHLGLADPQPWDLSYAGEKLREAKYAFSETEVKKYFPVVGK					
m128-1	ARRAKPYAEKDIAEVKAFARESLNLADLPQWDLYASEKLREAKYAFSETEVKKYFPVVGK					
	310	320	330	340	350	360
g128-1.pep	370	380	390	400	410	420
	VLAGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQONGKTIGGVYMDLYAREGKRGGAWM					
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQONGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
g128-1.pep	430	440	450	460	470	480
	NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILTLFHETGHGLHLLTQVD					
	430	440	450	460	470	480
g128-1.pep	490					
	ELGVSGINGVK					
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 491>:

a128-1.seq

```
1  ATGACTGACA ACGCACTGCT CCATTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATGCGCGAAG
101 CGCGCGAACA AATCGCGGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTGC AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTG
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
251 CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAACTC AACACGATC
401 TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCCG ACAGCAGGCA
451 GAATTGGCAA AACTGCAAA CAAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACGTACCT GCCGTCATCC AATACGCGCA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCTACGTT ACCCGCGCCA GCGAGTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCGCGCGCG CCAAACCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCGCG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCCG
1101 CCAAATCAAA AAACCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGCGGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
```



```
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAAC TCCAGTCAG
1501 TTTATGAAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAACCTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCT
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCGC CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 CGGCGAGAAT CCTTCAAAGC CTCCGCGGA CGGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA
```

This corresponds to the amino acid sequence <SEQ ID 492; ORF 128-1.a>:

a128-1.pep

```
1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWGA
51 NTVEPLTGIT ERVGRIVGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLPQ WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDARELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLEKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRWFQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAA*
```

m128-1/a128-1 97.8% identity in 677 aa overlap

```
10 20 30 40 50 60
a128-1.pep MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
|||||
m128-1 MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
10 20 30 40 50 60

70 80 90 100 110 120
a128-1.pep ERVGRIVGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
|||||
m128-1 ERVGRIVGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
70 80 90 100 110 120

130 140 150 160 170 180
a128-1.pep TLSHAQKTKLNHDRLDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
|||||
m128-1 TLSPAQKTKLNHDRLDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
130 140 150 160 170 180

190 200 210 220 230 240
a128-1.pep FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
|||||
m128-1 FDDAAPLAGIPEDALAMFAAAAQSESKTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
190 200 210 220 230 240

250 260 270 280 290 300
a128-1.pep TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
|||||
m128-1 TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
250 260 270 280 290 300

310 320 330 340 350 360
a128-1.pep ARRAKPYAEKDLAEVKAFARESLGLADLPQWDLGYAGEKLREAKYAFSETEVKKYFPVGK
|||||
```



Query: 363 NGLFAQIKKLYGIGFTE-KTVPVWHKDVRYPFEL-QQNGETIGGVYMDLYAREGKRGGAWM 420  
 +GLF IK+++ I E K V WHKDVR+F+L +N + G Y+DLYARE KRGGAWM  
 Sbjct: 365 SGLFELIKRIFNIRAVERKGVDTWHKDVRFFDLIDENDQLRGSFYLDLYAREHKRGGAWM 424

Query: 421 NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEIXXXXXXXXXXXXXXXXXXQVD 480  
 +D GR+R DG+++ P AYL CNF P+G K A +H+E+ Q+D  
 Sbjct: 425 DDCIGRKRKLDGSIETPVAYLTCNFNAPIGNKPALFTHNEVTTLFHEFGHGIHMLTQID 484

Query: 481 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ 540  
 V+GINGV WDAVELPSQFMEN+ WE LA +S H ETG PLPKE ++L AKNFQ  
 Sbjct: 485 VSDVAGINGVPWDVELPSQFMENWCWEEELAFISGHYETGEPLPKELTQLLKAQNFQ 544

Query: 541 RGMFLVRQMEFALFDMMIYSEDDEGRKKNWQVLDVSRKEVAVVRPPEYNRFANSFGHIF 600  
 MF++RQ+EF +FD ++ D + L SV+ +VAV++ ++ R +SF HIF  
 Sbjct: 545 AAMFILRQLFEGFIFDFRLHHTFDAETNQILDTLKSVKSQVAVIKGVWDWARAPHSFHIF 604

Query: 601 XXXXXXXXXXXXWAEVLSADAYAAFEESDDV-AATGKRFWQEILAVGGSRSAAESFKAFR 659  
 WAEVLSADAY+ FEE TKG F EIL GGS E FK FR  
 Sbjct: 605 AGGYAAGYYSYLWAEVLSADAYSRFEEGIFNPITGKSFLDEILTRGGSEEPMELFKRFR 664

Query: 660 GREPSIDALLRHSGFDN 676  
 GREP +DALLRH G N  
 Sbjct: 665 GREPQLDALLRHKGIMN 681

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 493>:

g129.seq  
 1 ATGCTTTCAC CTCCTCGGCG TAAACGGCG GCACATCAAT CAAGCCGTCT  
 51 TTCATTTGCG TCGGAAAAA ATGCGGCGTG TTGCCGTGAT CAAAATCAAT  
 101 ATCGTGACG ATCCAGCCCA AATCGCGGTT TGCCTCGCTT TCCGATAACG  
 151 CCGACGCGCG CAGCGGTTCA CCCTTATCCG CGCTTTCGCC ATTTGCCCTT  
 201 TCAGGCTGCG GGCATAGGGG CGGAACAGGC GCGGTCGAA TCCTGTTTCA  
 251 TCCGACAAA CGCGTTGGCA GTCGAAAAT CCGGCCGGCC GTGTCAAATA  
 301 ATGCGTTACT TTGGCCGGGT CTTGTCCTTT GTAAGCGGCG GTCTTTTTTT  
 351 GCGCGCCATC CGCATCTGTT TGGGCGCATG GCAAACGGCG GCTGCCGTAC  
 401 AATCAAAATG TTTGGCGATT TCATGCAGAC AGGCATCCGG ATGCCGCCCG  
 451 ACATATCGAG CCGGTTTTTG CCTATCCGAT TTGGCGGCAT TTAGGCCGGT  
 501 AACTGA

This corresponds to the amino acid sequence <SEQ ID 494; ORF 129.ng>:

g129.pep  
 1 MLSPRRKTA AHQSSRLSFA CGKNAACCRD QNQYRAASSP NRGLPRFPIT  
 51 PTAAAVHPYP RFRHLPFQAA GIGAEQAAVE SCFIRTNALA VGKSGRPCQI  
 101 MRYFGRVLSF VSGGLFLRAI RICLGAWQTA AAVQSKCLAI SCRQASGCRP  
 151 TYRAGFCLSD LAAFRPVT\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 495>:

m129.seq (partial)  
 1 ..TATCTGCGCT TTCACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA  
 51 ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG  
 101 GAAAATTCGG CCGGCTGTGT CAAATAATGC GTTACTTTGG CCGGGTCTTG  
 151 TTCTTTGTAA GTGGTGGTCT TTTTTTGCGC GTTATCCCCA TCTGTTTGAG  
 201 TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTCAT  
 251 GCAGATAGGC ATCCGGGTGT TGCCCAACAT ATTGAGCCGG TTTTGCCTA  
 301 TCCGATTGA CGGCATTAG ACCGGTAACT TGA

This corresponds to the amino acid sequence <SEQ ID 496; ORF 129>:

m129.pep (partial)  
 1 ..YLRFHLPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGRLC QIMRYFGRVL  
 51 FVSGGLFLR VIPICLSAXQ MVAQVSKCL AISCRXASGC CPTYXAGFCL  
 101 SDLTAFRPVT \*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 129 shows 79.1% identity over a 110 aa overlap with a predicted ORF (ORF 129.ng) from *N. gonorrhoeae*:

m129/g129

```

m129.pep                               10      20      30
                                         YLRFHYLPFQAAGIGTEQVAVKSCFIQINT
g129      RDQNQYRAASSPNRGLPRFPITPTAAAVHPYPRFRHLPPFQAAGIGAEQAAVESCFIRTNA
           30      40      50      60      70      80

m129.pep                               40      50      60      70      80      90
           LVVGKFGRLCQIMRYFGRVLFVSGGLFLRVIPICLSAXQMVAAVQSKCLAISCRXASGC
g129      LAVGKSGRPCQIMRYFGRVLSFVSGGLFLRAIRICLGAWQTAAAVQSKCLAISCRQASGC
           90     100     110     120     130     140

m129.pep                               100     110
           CPTYXAGFCLSDLTAFRPVTX
g129      RPTYRAGFCLSDLAAPFRPVTX
           150     160

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 497>:

```

a129.seq (partial)
1   TATCTGCGCT TTCACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA
51  ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG
101 GAAAATTCGG CCAGCTGTGT CAAATAATGC GTTACTTTGG CCGGCTCTTG
151 TTCTTTGTAA GTGGTGGTCT TTTTTCGCGC GTTATCCCCA TCTGTTTGAG
201 TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTCAT
251 GCAGATAGGC ATCCTGGTGT TGCCCAACAT ATTGAGCCGG TTTTGCCTA
301 TCCGATTGTA CGGCATTTAG ACCGTAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 498; ORF 129.a>:

```

a129.pep (partial)
1   YLRFHYLPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGQLC QIMRYFGRVL
51  FFVSGGLFLR VIPICLSA*Q MVAAVQSKCL AISCR*ASWC CPTY*AGFCL
101 SDLTAFRPVT *

```

**m129/a129 98.2% identity in 110 aa overlap**

```

m129.pep                               10      20      30      40      50      60
           YLRFHYLPFQAAGIGTEQVAVKSCFIQINTLVVGKFGRLCQIMRYFGRVLFVSGGLFLR
a129      YLRFHYLPFQAAGIGTEQVAVKSCFIQINTLVVGKFGQLCQIMRYFGRVLFVSGGLFLR
           10      20      30      40      50      60

m129.pep                               70      80      90      100     110
           VIPICLSAXQMVAAVQSKCLAISCRXASGCCPTYXAGFCLSDLTAFRPVTX
a129      VIPICLSAXQMVAAVQSKCLAISCRXASWCCPTYXAGFCLSDLTAFRPVTX
           70      80      90      100     110

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 499>:

```

g130.seq
1   ATGAAACAAC TCCGCGACAA CAAAGCCCAA GGCTCTGCAC TGTTTACCTT
51  TGTGAGCGGT ATCGTTATTG TTATTGCAGT CCTTTATTTT CTGATTAAGC
101 TGGCGGGCAG TGGATCGTTC GGCGATGTCG ATGCCACTAC GGAAGCGGCA
151 ACGCAGACCC GCATCCAGCC TGTCCGACAA TTGACGATGG GTGACGGCAT
201 CCCCGTCGGC GAACGCCAAG GCGAACAGAT TTCCGGCAAA ATCTGTATCC

```

```

251 AATGCCACGC GGC GGACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301 AACGGCGACT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCCA
351 ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGCAG
401 ACCTGACCGA TCAGGAACTC AAACGGGCGA TTACCTACAT GGCGAATAAA
451 AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501 CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG
551 CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601 AAAAAAGTCT TCGAAGCAAC CTGTCAGGTG TGCCACGGCG GTTCGATTCC
651 CGGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701 AAGGCAAAGA AACCTTGAC ACATGCCCC TTGAAGGCTT TAACGCGATG
751 CCGGCCAAAG GCGGCAATGC AGGTTTGAGC GATGACGAAG TCAAAGCGGC
801 TGTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 500; ORF 130.ng>:

```

g130.pep
  1 MKQLRDNKAQ GSALFTLVSG IVIVIAVLYF LIKLAGSGSF GDVDATTEAA
 51 TQTRIOPVGQ LTMGDGIPVG ERQGEQIFGK ICIQCHAADS NVPNAPKLEH
101 NGDWAPRIAQ GFDTLFQHAL NGFNAMPAKG GAADLTDQEL KRAITYMANK
151 SGGSFPPNDE AAPADNAASG TASAPADSAA PAEAKAEDKG AAPAVGVVDG
201 KKVFEATCQV CHGGSIPGIP GIGKDDWAP RIKKGKETLH KHALEGFNAM
251 PAKGGNAGLS DDEVKAAVDY MANQSGAKF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 501>:

```

m130.seq (partial)
  1 ..GGCGAACAGA TTTTCGGCAA AATCTGTATC CAATGCCACG CGGCGGACAG
 51 CAATGTGCCG AACGCTCCGA AACTGGAACA CAACGGCGAT TrGGCACC GC
101 GTATCGGCAA GGCTTCGATA CCTTGTTCCA ACACGCGCTG AACGGCTTTA
151 ACGCCATGCC TGCAAAAGGC GGTGCGGCAG ACCTGACCGA TCAGGAACTT
201 AAACGGGCGA TTA CTTACAT GGC GAACAAA AGCGGCGGTT CTTTCCCGAA
251 TCCTGATGAG GCTGCGCCTG CCGACAATGC CGCTTCAGGA ACAGCTTCTG
301 CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG CGAAGGCAGA AGACAAGGGT
351 GCGGCACCCC TGCGGTCGGC GTTGACGGTA AAAAAGTCTT CGAAGCAACC
401 TGTCAGGTGT GCCACGGCGG TTCGATTCCC GGTATTCCCG GCATAGGCAA
451 AAAAGACGAT TGGGCACCGC GTATCAAAAA AGGCAAAGAA ACCTTGACACA
501 AACACGCCCT TGAAGGCTTT AACGCGATGC CTGCCAAArG CGCAATGCA
551 GGT TTGAGCG ATGACGAAGT CAAAGCGGCT GTTGACTATA TGGCAAACCA
601 ATCCGGTGCA AAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 502; ORF 130>:

```

m130.pep (partial)
  1 ..GEQIFGKICI QCHAADSNVP NAPKLEHNGD XAPRIQGFDT LFQHALNGFN
 51 AMPAKGGAAD LTDQELKRAI TYMANKSGGS FPNPDEAAPA DNAASGTASA
101 PADSAPAEA KAEDKGAAPA VGVDGKKVFE ATCQVCHGGS IPGIPGIGKK
151 DDWAPRIKKG KETLHKHALE GFNAMPAXG NAGLSDDDEVK AAVDYMANQS
201 GAKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 130 shows 98.1% identity over a 206 aa overlap with a predicted ORF (ORF 130.ng) from *N. gonorrhoeae*:

```

m130/g130
m130.pep
                                     10      20      30
                                     GEQIFGKICIQCHAADSNVPNAPKLEHNGD
                                     |||
g130      DATTEAATQTRIOPVGQLTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
                                     50      60      70      80      90     100

                                     40      50      60      70      80      89
m130.pep      XAPRI-QGFDTLFQHALNGFNAMPAGGAADLTDQELKRAITYMANKSGGSFPNPDEAAP
                                     |||
g130      WAPRIAQGFDTLFQHALNGFNAMPAGGAADLTDQELKRAITYMANKSGGSFPNPDEAAP
                                     110     120     130     140     150     160

```

```

          90      100      110      120      130      140
m130.pep  ADNAASGTASAPADSAAPAEAKAEDKGAA-PAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          |||||
g130      ADNAASGTASAPADSAAPAEAKAEDKGAAAPAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          170      180      190      200      210      220

          150      160      170      180      190      200
m130.pep  KKDDWAPRIKKGKETLHKHALEGFNAMPKXGNAGLSDDEVKAAVDYMANQSGAKFX
          |||||
g130      KKDDWAPRIKKGKETLHKHALEGFNAMPKGGNAGLSDDEVKAAVDYMANQSGAKFX
          230      240      250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 503>:

```

a130.seq
1  ATGAAACAAC TCCGCGACAA CAAAGCCCAA GGCTCTGCAC TGTTTACCTT
51  TGTGAGCGGT ATCGTTATTG TTATTGCACT CCTTTATTTT CTGATTAAGC
101 TGGCGGGCAG CGGCTCGTTC GGCGATGTCG ATGCCACTAC GGAAGCAGCA
151 ACGCAGACCC GTATCCAGCC TGTCGGACAA TTGACGATGG GCGACGGCAT
201 CCCCCTCGGC GAACGCCAAG GCGAACAGAT TTTCGGGCAA ATCTGTATCC
251 AATGCCACGC GCGCGACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301 AACGGCGATT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCCTA
351 ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGTAG
401 ACCTGACCGA TCAGGAATC AAACGGGCGA TTAATTACAT GCGCAACAAA
451 AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501 CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGACGCT CCGGCAGAAG
551 CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601 AAAAAAGTCT TCGAAGCAAC CTGTCAGGTG TGCCACGGCG GTTCGATTCC
651 CCGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701 AAGGCAAAGA AACCTTGAC AAACACGCCC TTGAAGGCTT TAACGCGATG
751 CCTGCCAAAG GCGGCAATGC AGGTTTGAGC GATACGAAG TCAAAGCGGC
801 TGTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 504; ORF 130.a>:

```

a130.pep
1  MKQLRDNKAQ GSAFLTLVSG IVIVIAVLYF LIKLAGSGSF GDVDATTEAA
51  TQTRIQPVGQ LTMGDGIPVG ERQGEQIFGK ICIQCHAADS NVPNAPKLEH
101 NGDWAPRIAQ GFDTLFQHAL NGFNAMPKAG GAVDLTDQEL KRAITYMANK
151 SGGSFNPDE AAPADNAASG TASAPADSAA PAEAKAEDKG AAAPAVGVDG
201 KKVFEATCQV CHGGSIPGIP GIGKKDDWAP RIKKGKETLH KHALEGFNAM
251 PAKGGNAGLS DDEVKAAVDY MANQSGAKF*

```

m130/a130 97.6% identity in 206 aa overlap

```

          10      20      30
m130.pep  GEQIFGKICIQCHAADSNVPNAPKLEHNGD
          |||||
a130      DATTEAATQTRIQPVGQLTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
          50      60      70      80      90      100

          40      50      60      70      80      89
m130.pep  XAPRI-QGFDTLFQHALNGFNAMPKGGAADLTDQELKRAITYMANKSGGSFNPDEAAP
          |||||
a130      WAPRIAQGFDTLFQHALNGFNAMPKGGAVDLTDQELKRAITYMANKSGGSFNPDEAAP
          110      120      130      140      150      160

          90      100      110      120      130      140
m130.pep  ADNAASGTASAPADSAAAEAKAEDKGAA-PAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          |||||
a130      ADNAASGTASAPADSAAAEAKAEDKGAAAPAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          170      180      190      200      210      220

          150      160      170      180      190      200
m130.pep  KKDDWAPRIKKGKETLHKHALEGFNAMPKXGNAGLSDDEVKAAVDYMANQSGAKFX

```

```

|||||
a130  KKDDWAPRIKKGKETLHKHALEGFNAMPAKGGNAGLSDDEVKAAVDYMANQSGAKFX
      230      240      250      260      270      280

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 505>:

```

g132.seq
  1  ATGGAAGCCT TCAAAACCCT AATTGGATT ATTAATATTA TTTCCGCTTT
 51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101  GCGCGACCTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151  GCCGGCAACG CCAACTTcct CAgccGCTCG AccGccGTTG CAGCAACatt
201  tttctttTGca acctgcAtgg gctatggTgt atattcacac CCACACGACA
251  AAACACGGTT TGGACTtcag caacataCGA CAGACTCAGC AagcACCCAA
301  ACCcgtAAGC AATACCGAAC CTTCTGCCCC TGTTCTCTAG CAGCAGAAAT
351  AACagtTTTT CAAATgcccga catGgtga

```

This corresponds to the amino acid sequence <SEQ ID 506; ORF 132.ng>:

```

g132.pep
  1  MEAFKTLIWI INIISALAVI VLVLQHGKG ADAGATFGSG SGSAQGVFGS
 51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QHTTDSASTQ
101  TRKQYRTFCP CSSAAEITVF QMPTW*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 507>:

```

m132.seq (partial)
  1  ATGGAACCCT TCAAAACCTT AATTGGATT GTTAATTTAA TTTCCGCTTT
 51  GGCCGTCCTT GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101  GCGCGACTTT CGGA...

```

This corresponds to the amino acid sequence <SEQ ID 508; ORF 132>:

```

m132.pep (partial)
  1  MEPFKTLIWI VNLISALAVF VLVLQHGKG ADAGATFG...

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 132 shows 89.5% identity over a 38 aa overlap with a predicted ORF (ORF 132.ng) from *N. gonorrhoeae*:

```

m132/g132
      10      20      30
m132.pep  MEPFKTLIWI VNLISALAVF VLVLQHGKG ADAGATFG
          || |||||:|:|||||:|||||
g132      MEAFKTLIWI INIISALAVI VLVLQHGKG ADAGATFGSGSGSAQGVFGSAGNANFLSRS
          10      20      30      40      50      60

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 509>:

```

a132.seq
  1  ATGGAAGCCT TCAAAACCCT AATTGGATT GTTAATATAA TTTCCGCTTT
 51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101  GCGCGACTTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151  GCCGGCAACG CTAACCTCCT CAGCCGCTCG ACCGCCGTTG CAGCAACATT
201  TTTCTTTGCA ACCTGCATGg GCTATGGTGT ATATTACAC CCACACGACA
251  AAACACGGTT TGGACTTCAG CAACGTACAA CAAACTCAGC AAGCACCCAA
301  ACCCGTAAGC AATACCGAAC CTTCTGCCCC TGTTCTCTAG CAGCAGAAAT
351  AACAGTTTTT CAAATGCCGA CATGGTGA

```

This corresponds to the amino acid sequence <SEQ ID 510; ORF 132.a>:

```

a132.pep
  1  MEAFKTLIWI VNIISALAVI VLVLQHGKG ADAGATFGSG SGSAQGVFGS
 51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QRTTNSASTQ
101  TRKQYRTFCP CSSAAEITVF QMPTW*

```

m132/a132 92.1% identity in 38 aa overlap

	10	20	30	
m132.pep	MEPFKTLIWIIVNLISALAVFVLVLLQHGKGADAGATFG			
	:     :     :			
a132	MEAFKTLIWIIVNIISALAVIVLVLLQHGKGADAGATFGSGSGSAQGVFGSAGNANFLSRS			
	10	20	30	40 50 60

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 511>:

g134.seq

1	ATGTCCCAAG	AAATCCTCGA	CCAAGTGC	CGCCGCCGCA	CGTTTGCCAT
51	CATCTCCAC	CCC	GATGCGG	GTA	AAACCAC
101	TGTTTTCGGG	CGCGATTCAA	AGCGCAGGCA	CGGTGAAAGG	TAAGAAAACC
151	GGCAAATTCG	CCACCTCCGA	CTGGATGGAC	ATCGAGAAGC	AGCGCGGCAT
201	TTCCGTGGCA	TCAAGCGTGA	TGCAGTTCGA	CTACAAAGAC	CACACCGTCA
251	ACCTCTTGGA	CACGCCGGGA	CACCAAGACT	TCTCCGAAGA	CACCTACCGC
301	GTTTTAACCG	CAGTGGACAG	CGCCTTGATG	GTCATCGACG	CGGCAAAAGG
351	CGTGGAAAGCG	CAAACCATCA	AACTCTTGAA	CGTCTGCCGC	CTGCGCGATA
401	CGCCGATTGT	TACCTTCATG	AACAAATACG	ACCGCGAAGT	GCGCGATTCT
451	TTGGAACCTCT	TGGACGAAGT	GGAAGACATC	CTGCAAATCC	GCTGCGCGCC
501	CGTTACCTGG	CCGATCGGTA	TGGGCAAAAA	CTCAAGGGC	GTGTACCACA
551	TCCTGAACGA	CGAAATCTAT	CTCTTTGAAG	CGGGCGGCGA	ACGCCTGCCG
601	CACGAGTTCG	ACATCATCAA	AGGCATAAAC	AATCCCGAAT	TGGAACAACG
651	CTTTCCGTTG	GAAATCCAGC	AGTTGCGCGA	CGAAATCGAA	TTGGTGCAGG
701	CGGCTTCCAA	CGAATTTAAT	CTCGacgaAT	TTCTCGccgG	CGAACTCACG
751	CCAGTGTCT	TCGGCTCTGC	GATTAACAAC	TTCGGCATT	AGGAAATCCT
801	CAATTCAATG	ATTGACTGGG	CACCCGCACC	GAAACCGCGC	GACGCGACCA
851	TGCGCATGGT	CGGGCCGGAC	GAGCCGAAAT	TTTCCGGATT	TATCTTTAAA
901	ATCCAAGCCA	ATATGGACCC	GAAACACCGC	GACCGTATCG	CCTTCTTGCG
951	CGTCTGCTCC	GGTAAATTG	AGCGCGGCAT	GAAGATGAAA	CACCTGCGTA
1001	TCAACCGCGA	AATCGCCGCC	TCCAGCGTAG	TAACCTTCAT	GTCGCACGAC
1051	CGCGAACTGG	CGGAAGAAGC	CTACGCCGGC	GACATCATCG	GCATCCCGAA
1101	CCACGGCAAC	ATCCAAATCG	GCGACAGCTT	CTCCGAAGGC	GAACAACTGG
1151	CGTTTACCGG	CATCCCATTC	TTCCGCGCCG	AACTGTTCCG	CAGCGTCCGC
1201	ATCAAAAACC	CGCTGAAAAT	CAAACAATG	CAAAAAGGTT	TGCAACAAC
1251	CGCGGAAGAA	GGTGC	GGTTC	AAGTATTCAA	ACCGATGAGC
1301	TGATTTGGG	TGCGGTGCGC	GTGTTGCACT	TTGAAGTCGT	AACCTCACGC
1351	CTCGCCAACG	AATACGGCGT	GGAAGCCGTG	TTCGACAGCG	CATCCATCTG
1401	GTCGCGCGCG	TGGGTATCGT	GCGACGACAA	GAAAAAACTG	GCGGAATTTG
1451	AAAAAGCCAA	CGCAGGCAAC	CTCGCCATCG	ACGCAGGCGG	CAACCTCGCC
1501	TACCTCGCCC	CCAACCGCGT	GAATTTGGGG	TTGACGCAAG	AACGCTGGCC
1551	GGACATCGTG	TTCCACGAAA	CGCGCGAACA	TTCGGTCAAA	CTCTAA

This corresponds to the amino acid sequence <SEQ ID 512; ORF 134.ng>:

g134.pep

1	MSQEILDQVR	RRRTFAISH	PDAGKTLTE	KLLLFSGAIQ	SAGTVKGKKT
51	GKFATSDWMD	IEKQRGISVA	SSVMQFDYKD	HTVNLLDTPG	HQDFSEDITYR
101	VLTAVDSALM	VIDAAKGVEA	QTIKLLNVCR	LRDTPIVTFM	NKYDREVRDS
151	LELLDEVEDI	LQIRCAPVTW	PIGMGKNFKG	VYHILNDEIY	LFEAGGERLP
201	HEFDIIKGIN	NPELEQRFPL	EIQQLRDEIE	LVQAASNEFN	LDEFLAGELT
251	PVFFGSAINN	FGIQEILNSL	IDWAPAPKPR	DATMRMVGPD	EPKFSGFIFK
301	IQANMDPKHR	DRIAFLRVCS	GKFERGMKMK	HLRINREIAA	SSVVTFMSHD
351	RELAEEAYAG	DIIGIPNHGN	IQIGDSFSEG	EQLAFTGIPF	FAPELFRSVR
401	IKNPLKIKQL	QKGLQQLGEE	GAVQVFKPMS	GADLILGAVG	VLQFEVVTSR
451	LANEYGVAV	FDSASIWSAR	WVSCDDKKKL	AEFEKANAGN	LAIDAGGNLA
501	YLAPNRVNLG	LTQERWPDIV	FHETREHSVK	L*	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 513>:

m134.seq

1	ATGTCCCAAG	AAATCCTCGA	CCAAGTGC	CGCCGCCGCA	CGTTTGCCAT
51	CATCTCCAC	CCTGACGCAG	GTA	AAACCAC	GTTGACTGAA
101	TGTTTTCGGG	CGCGATTGAG	AGCGCGGGTA	CGGTAAAAGG	CAAGAAAACC
151	GGCAAATTCG	CCACTTCCGA	CTGGATGGAA	ATCGAGAAGC	AGCGCGGCAT



```

201 TTCCGTGGCA TCAAGTGTGA TGCAGTTCGA TTACAAAGAC CACACCGTCA
251 ACCTCTTGGA CACGCCGGGA CACCAAGACT TC'TCCGAAGA CACCTACCGC
301 GTTTTAACCG CCGTGGACAG CGCATTAAATG GTCATCGACG CGGCAAAAGG
351 CGTGGAAGCG CAAACCATCA AGCTCTTAAA CGTCTGCCGC CTGCGCGATA
401 CACCGATTGT TACGTTTATG AACAAATACG ACCGCGAAGT GCGCGATTCC
451 CTGGAACCTT TGGACGAAGT GGAAAACATT TTAAAAATCC GCTGCGCGCC
501 CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAAGGGC GTGTACCACA
551 TCCTGAACGA TGAAATTTAT CTCTTTGAAG CTGGCGGCGA ACGCCTGCCG
601 CACGAGTTCG ACATCATCAA AGGCATCGAT AATCCTGAAT TGGAACAACG
651 CTTTCCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701 CGCGATTCAA CGAGTTTAAAT CTCGACGAAT TCCTCGCCGG CGAATCCACG
751 CCCGTATTCT TCGGCTCTGC GATTAACAAC TTCGGTATTC AGGAAATCCT
801 CAATTCAATG ATTGACTGGG CGCCCGCGCC GAAACCGCGC GACGCGACCG
851 TACGTATGGT CGAGCCGGAC GAGCCGAAGT TTTCCGGATT TATCTTCAAA
901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATTG CCTTCTTGCG
951 CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAGATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TTACCTTCAT GTCGCACGAC
1051 CGCGAGCTGG TTGAAGAAGC CTACGCCGGC GACATTATCG GCATCCCGAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACTGG
1151 CGTTCACCGG CATCCCATTG TTCGCACCCG AACTGTTCCG CAGCGTACGC
1201 ATCAAAAACC CGCTGAAAAT CAAACAACG CAAAAGGCT TGCAACAGCT
1251 CGGCGAAGAA GGCGCGGTGC AGGTGTTCAA ACCGATGAGC GGCGCGGATT
1301 TGATTTTGGG CGCGGTGCGC GTGTTGCAGT TTGAAGTCGT TACCTCGCGC
1351 CTCGCCAACG AATACGGCGT AGAAGCCGTG TTCGACAGCG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAACTG GCTGAATTTG
1451 AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC
1501 TACCTCGCCC CCAACCGCGT GAATTTGGGA CTCACGCAAG AACGTTGGCC
1551 GGACATCGTG TTCCACGAAA CACGCGAACA TTCGGTCAAA CTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 514; ORF 134>:

```

m134.pep
1  MSQEILDQVR RRRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
51  GKFDATSDWME IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101 VLTAVDSALM VIDAAKGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS
151 LELLDEVENI LKIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGID NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATVRMVEPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTMMSHD
351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451 LANEYGVEAV FDSASIWSAR WVSCDDKKKL AEFKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 134 shows 98.7% identity over a 531 aa overlap with a predicted ORF (ORF 134.ng) from *N. gonorrhoeae*:

m134/g134

	10	20	30	40	50	60
m134.pep	MSQEILDQVRRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME					
g134	MSQEILDQVRRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD					
	10	20	30	40	50	60
	70	80	90	100	110	120
m134.pep	IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAKGVEA					
g134	IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAKGVEA					
	70	80	90	100	110	120
	130	140	150	160	170	180

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

134

m134.pep	QTIKLLNVCRLRDTPIVTFMKNKYDREVRDSLELLDEVENILKIRCAPVTWPIGMGNFKG
g134	QTIKLLNVCRLRDTPIVTFMKNKYDREVRDSLELLDEVEDILQIRCAPVTWPIGMGNFKG
	130 140 150 160 170 180
m134.pep	VYHILNDEIYLF EAGGERLPHEFDIIKGINNPELEQRFPLEIQQLRDEIELVQAASNEFN
g134	VYHILNDEIYLF EAGGERLPHEFDIIKGINNPELEQRFPLEIQQLRDEIELVQAASNEFN
	190 200 210 220 230 240
m134.pep	LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATVRMVEPDEPKFSGFIFK
g134	LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATMRMVGPDPEPKFSGFIFK
	250 260 270 280 290 300
m134.pep	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFM SHDRELVEEAYAG
g134	IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFM SHDRELVEEAYAG
	310 320 330 340 350 360
m134.pep	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE
g134	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE
	370 380 390 400 410 420
m134.pep	GAVQVFKPMMSGADLILGAVGVQLQFEVVT SRLANEYGV EAVFDSASIWSARWVSCDDKKKL
g134	GAVQVFKPMMSGADLILGAVGVQLQFEVVT SRLANEYGV EAVFDSASIWSARWVSCDDKKKL
	430 440 450 460 470 480
m134.pep	AEFEKANAGNL AIDAGGNLAYLAPNVRNLGLTQERWPDIVFHETREHSVKLX
g134	AEFEKANAGNL AIDAGGNLAYLAPNVRNLGLTQERWPDIVFHETREHSVKLX
	490 500 510 520 530

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 515>:

a134.seq

1	ATGTCCCAAG	AAATCCTCGA	CCAAGTGCGC	CGCCGCCGCA	CGTTTGCCAT
51	CATCTCCAC	CCTGACGCAG	GTAAAACCAC	GTTGACTGAA	AAACTCTTGC
101	TGTTTTCAGG	TGCGATTCAA	AGCGCGGGTA	CGGTAAGG	CAAGAAAACC
151	GGCAAATTCG	CCACCTCCGA	CTGGATGGAC	ATCGAGAAGC	AGCGCGGCAT
201	TTCCGTGGCA	TCAAGCGTGA	TGCAGTTCGA	CTATAAAGAC	CACACCGTCA
251	ACCTTTTGA	CACGCCGGA	CACCAAGACT	TCTCCGAAGA	CACCTACCGC
301	GTTTGTACCG	CCGTCGATAG	TGCCTTGATG	GTCATCGACG	CGGCAAAAGG
351	CGTGGAAGCG	CAAACCATCA	AACTCTTGAA	CGTCTGCCGC	CTGCGCAATA
401	CGCCGATTGT	TACGTTTCATG	AACAAATACG	ACCGCGAAGT	GCGCGATTCC
451	CTGGAATTGC	TGGACGAAGT	GGAACATC	CTGCAAATCC	GCTGCGCGCC
501	CGTAACCTGG	CCGATCGGCA	TGGGCAAAA	CTTCAAAGGC	GTGTACCACA
551	TCCTGAACGA	CGAAATCTAT	CTCTTTGAAG	CGGGCGGCGA	ACGCTTGCCG
601	CACGAGTTCG	ACATCATCAA	AGGCATCGAT	AATCCCGAAT	TGGAACAACG
651	CTTCCGTTA	GAAATACAGC	AGTTGCGCGA	CGAAATCGAA	TTGGTGCAGG
701	CGGCTTCCAA	CGAGTTCAAT	CTCGACGAAT	TCCTCGCCGG	CGAACTCAGC
751	CCCGTATTCT	TCGGCTCTGC	GATTAACAAC	TTCCGTATT	AGGAAATCCT
801	CAATTCATTG	ATTGAATGGG	CGCCCGCGCC	GAAACCACGC	GATGCGACCG
851	TGCGTATGGT	CGAGCCGGAC	GAGCCGAAGT	TTTCCGATT	TATCTTCAA
901	ATCCAAGCCA	ATATGGACCC	GAAACACCGC	GACCGTATTG	CCTTCTTGCG

```

951 CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAAATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TAACCTTCAT GTCCCACGAC
1051 CGCGAGCTGG TTGAAGAAGC CTACGCCGGC GACATTATCG GTATCCCAAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACCTGA
1151 CGTTTACCGG CATCCCATTG TTCGCGCCCG AACTGTTCCG CAGCGTTCGC
1201 ATCAAAAACC CGCTGAAAAT CAAGCAACTG CAAAAAGGTT TGCAACAGCT
1251 TGGCGAAGAA GGTGCGGTGC AGGTGTTCAA ACCAATGAGC GGCGCGGATT
1301 TGATTTTGGG CGCGGTCGGC GTGTTGCACT TTGAAGTCGT TACCTCGCGC
1351 CTTGCCAACG AATACGGCGT GGAAGCCGTG TTCGACAACG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAAGT GCGGAATTG
1451 AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCGGGCGG CAACCTCGCC
1501 TACCTCGCCC CTAACCGCGT GAATCTGGGA CTCACGCAAG AACGCTGGCC
1551 GGACATCGTG TTCCACGAAA CGCGCGAGCA TTCGGTCAAA CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 516; ORF 134.a>:

```

a134.pep
1 MSQEILDQVR RRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
51 GKFATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101 VLTAVDSALM VIDAAKGVEA QTIKLLNVCR LRNTPIVTFM NKYDREVRDS
151 LELLDEVENI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGID NPELEQRFPL EIQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IEWAPAPKPR DATVRMVEPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLTFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVTSR
451 LANEYGVAV FDNASIWSAR WVSCDDKKKL AEFKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

m134/a134 98.9% identity in 531 aa overlap

```

m134.pep      10      20      30      40      50      60
MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
|||||
a134          10      20      30      40      50      60
MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD

m134.pep      70      80      90      100     110     120
IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDLSALMVIDAAKGVEA
|||||
a134          70      80      90      100     110     120
IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDLSALMVIDAAKGVEA

m134.pep     130     140     150     160     170     180
QTIKLLNVCRRLDTPIVTFMKNKYDREVRDSLELLDEVENILKIRCAPVTWPIMGKNFKG
|||||
a134         130     140     150     160     170     180
QTIKLLNVCRRLNTPIVTFMKNKYDREVRDSLELLDEVENILQIRCAPVTWPIMGKNFKG

m134.pep     190     200     210     220     230     240
VYHILNDEIYLFEEAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
|||||
a134         190     200     210     220     230     240
VYHILNDEIYLFEEAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN

m134.pep     250     260     270     280     290     300
LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATVRMVEPDEPKFSGFIFK
|||||
a134         250     260     270     280     290     300
LDEFLAGELTPVFFGSAINNFGIQEILNSLIEWAPAPKPRDATVRMVEPDEPKFSGFIFK

m134.pep     310     320     330     340     350     360
IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG
|||||
a134         310     320     330     340     350     360
IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG

```

	370	380	390	400	410	420
m134.pep	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIKQLKGLQQLGEE					
a134	DIIGIPNHGNIQIGDSFSEGEQLTFTGIPFFAPELFRSVRIKNPLKIKQLKGLQQLGEE					
	370	380	390	400	410	420
	430	440	450	460	470	480
m134.pep	GAVQVFKPMSGADLILGAVGVQLQFEVVTSRLANEYGVFAVFDNASIWSARWVSCDDKKKL					
a134	GAVQVFKPMSGADLILGAVGVQLQFEVVTSRLANEYGVFAVFDNASIWSARWVSCDDKKKL					
	430	440	450	460	470	480
	490	500	510	520	530	
m134.pep	AEFEKANAGNLAI DAGGNLAYLAPNRVNLGLTQERWPDIVFHETREHSVKLX					
a134	AEFEKANAGNLAI DAGGNLAYLAPNRVNLGLTQERWPDIVFHETREHSVKLX					
	490	500	510	520	530	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 517>:

```

g135.seq
1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCG
51  TTCGGACgGC AGCCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCTG CAGGGTTCGG CGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GTGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCGATTCCC ATCATCAATG
401 AAAACGACAC GGTTCGCTT GAGGAGTTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT
501 GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CGGGTTCGGC AAACGGCAGC GGCGGTATGC TGACCAAAAT
651 CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
701 CCTCACTCAA ACCCGATTCA TTGGCCGAAG CCGCCGAACA TCAGGCGGAC
751 GGCTCGTTTT TCGTcccCcg tgCCAAAGGT TTGCGGACAC AGAAGCAATG
801 GctggCGTTC TATTcggaaa gcgggGgag cgttTatgtg gacgaaagtg
851 cggaacacgc tTgtccgaa caagggaaag cctgCTGA
  
```

This corresponds to the amino acid sequence <SEQ ID 518; ORF 135.ng>:

```

g135.pep
1  MKYKRIVFKV GTSSITRSDG SLRSGKIQTI TRQLAALHHA GHELVLVSSG
51  AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAAGAL SVLLQORRAIP IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHNHEIEM
201 AGGSGSANGT GGMLTKIKAA TIAAESGVPV YICSSLKPDS LAEAAEHQAD
251 GSFFVPRAKG LRTQKQWLAF YSESGGSVYV DESAEHALSE QGKAC*
  
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 519>:

```

m135.seq
1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
51  TTCGGACGGC AGTCTCTCGC GCGGCAAAAT CCAAACCATC ACCTGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCGG CAGGGTTCGG TGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCGCAAATC
301 CTGCTCAGCC GCGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCCC ATCATCAATG
401 AAAACGATAC GGTTCGCTT GAGGAATTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT
  
```

```

501 GACCGACATA GACGGTCTTT ACACGGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CGGGTTCGGC AAACGGCAGC GCGGGTATGC TGACCAAAAT
651 CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CTGCCGAACA TCAGGCGGAC
751 GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
801 GCTGGCGTTC TATTCCGAAA GCCGGGGCAG CGTTTATGTG GACGAAGGTG
851 CGGAACACGC TTTGTCCGAA CAGGGGAAAA GCCTGCTGAT GTCGGGCATT
901 GCCGGAATCG AAGGGCATTT TTCCCGTATG GACACCGTAA CCGTGTACAG
951 CAAGGCAACC AAACAGCCCC TGGGCAAAGG GCGCGTCCTG TTCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATCGCGTA AGGCGAAAGG CGTGTTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTTAG

```

This corresponds to the amino acid sequence <SEQ ID 520; ORF 135>:

```

m135.pep
  1 MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TCQLAALHHA GHELVLVSSG
 51 AVAAGFGALG FKKRPVKIAD KOASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAPV IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGlyTGNPNS NPDAVRLDKI EHINHEIEM
201 AGGSGSANGT GGMLTKIKAA TIAESGVPV YICSSLKPDA LAEAAEHQAD
251 GSFFVPRAKG LRTQKQWLAF YSESRSVYV DEGAEHALSE QGKSLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAEDLL KSRKAKGVFI
351 HRDDWISITP EIRLLLTEF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 135 shows 97.6% identity over a 294 aa overlap with a predicted ORF (ORF 135.ng) from *N. gonorrhoeae*:

m135/g135

```

              10      20      30      40      50      60
m135.pep      MKYKRIVFKVGTSSITHSDGSLSRGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g135           MKYKRIVFKVGTSSITRSDGSLSRGKIQTITRQLAALHHAGHELVLVSSGAVAAGFGALG
              10      20      30      40      50      60

              70      80      90     100     110     120
m135.pep      FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQIILSRADFADKRRYQNAGGAL
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g135           FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQIILSRADFADKRRYQNAGGAL
              70      80      90     100     110     120

              130     140     150     160     170     180
m135.pep      SVLLQRRAPVPIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGNPNS
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g135           SVLLQRRAPVPIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGNPNS
              130     140     150     160     170     180

              190     200     210     220     230     240
m135.pep      NPDAVRLDKIEHINHEIEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPDA
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g135           NPDAVRLDKIEHINHEIEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPD
              190     200     210     220     230     240

              250     260     270     280     290     300
m135.pep      LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESRSVYVDEGAEHALSEQGKSLMSGI
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g135           LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESRSVYVDESAEHALSEQGKACX
              250     260     270     280     290

              310     320     330     340     350     360

```

m135.pep AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFSGAAAEDLLKSRKAKGVFIHRDDWISITP

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 521>:

```
a135.seq
1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
51  TTCGGACGGC AGTCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCGG CAGGGTTTCGG TGCCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GCGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCCC ATCATCAATG
401 AAAACGATAC GGTTCGGTTC GAGGAATTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGTGCT
501 GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CGGGTTCGGC AAACGGCACA GGCGGTATGC TGAATAAAAT
651 CAAAGCGGCG ACGATTGCGA CCGAGTCCGG CGTACCGGTC TATATCTGTT
701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CGGCAGATAA TCAGGCGGAC
751 GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
801 GCTGGCGTTC TATTCGAAA GCAGGGGCGG CGTTTATGTG GACGAAGGTG
851 CGGAACACGC TTTGTCCGAA CAGGGA AAAA GCCTGCTGAT GTCGGGCATT
901 GCCGGAATCG AAGGGCATT TTTCCGTATG GACACCGTAA CCGTGTACAG
951 CAAGGCAACC AAACAGCCTT TGGGCAAAGG GCGAGTCCTG TTCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATTGCGTA AGGCGAAAGG CGTGTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTAG
```

This corresponds to the amino acid sequence <SEQ ID 522; ORF 135.a>:

```
a135.pep
1  MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TRQLAALHHA GHELVLVSSG
51  AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRRAVP IINENDTVSV EELKIGDNND
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRDLKI EHNHEIEM
201 AGGSGSANGT GGMLTKIKAA TIATESGVPV YICSSLKPDA LAEADNQAD
251 GSFFVPRAKG LRTQKQWLAF YSESRGGVYV DEGAEHALSE QGKSLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAAEDLL KLRKAKGVFI
351 HRDDWISITP EIRLLLTEF*
```

m135/a135 98.4% identity in 369 aa overlap

```

      10      20      30      40      50      60
m135.pep  MKYKRIVFKVGTSSITHSDGSLSRGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG
a135      MKYKRIVFKVGTSSITHSDGSLSRGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG
      10      20      30      40      50      60
m135.pep  FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLRADFADKRRYQNAGGAL
a135      FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLRADFADKRRYQNAGGAL
      70      80      90     100     110     120
m135.pep  SVLLQRRRAVPPIINENDTVSVEELKIGDNNDTSLAQVAAMIQADLLVLLTDIDGLYTGNPNS
a135      SVLLQRRRAVPPIINENDTVSVEELKIGDNNDTSLAQVAAMIQADLLVLLTDIDGLYTGNPNS
      130     140     150     160     170     180
m135.pep  NPDAVRDLKIEHNHEIEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPDA
a135      NPDAVRDLKIEHNHEIEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPDA
      190     200     210     220     230     240
```



g136.seq

This corresponds to the amino acid sequence <SEQ ID 524; ORF 136.ng>:

g136.pep

The following partial DNA sequence was identified in *N. meningitidis* <SEO ID 525>:

m136.seq

1	ATGGAACAA	ACGCTTCAAT	TCTTACCGCA	ACACGCCTTG	TATTTTCTGC
51	CGCTGCCGCA	CGGACAGGGA	TCGTTCTCTGC	CTGTTTTTTC	GCCTTCCCTG
101	CGGACGGTTT	GCGGTTTGTT	GATGACTGCC	TGCCAGTAGC	GGTAGATATC
151	CGCCAATGCA	TAAGGCAACT	CGGATTCCAG	TTCCGCAGC	TCGCCTTCTG
201	TGAATTGACG	ACGGATAGCG	CCGTTTTTCT	CTTCGTCGTA	AATACCGCCC
251	AATGCCATGA	TGGGATAAAA	CAACTCTTCA	AACGCTTCAT	CATCGACGGC
301	TTCAAACCAA	TCGGTCGGCA	CAATATCCAA	ACCGTAAAGA	TAAGCATTGC
351	ACCATTGTGA	AAAAATCGCTG	CCGCCGCTT	CGTTTTTCATA	CAGCCACAAA
401	TCGGGCAGTT	TTTTATCCGA	CATCGCGGTC	GTTGTTTCCA	TCGCCATTGC
451	CAAAACCAAG	CGTTCGATT	CGGAACGTT	GGTGGCCGTA	AATTGCGATT
501	CGTCGCCCAA	CACTTCGGGC	AGCCAGTCTGA	GCGGTGTCAA	TTTGTCCGGC
551	CCGCTCAACA	GCGCCGTCAT	AAAACCTTGA	ACCTCGTCGC	AACGCATCGT
601	GTTGCCTTGT	TCGCTTTTGG	CATCCAACAA	TTTCGCTCAAC	CGCCGTTTGG
651	ATGCTTCGGT	AAATTTTCGG	GAATCCATCA	TTTTCTTTTT	CAAATGGGTT
701	TTGCGCCCTA	TTATCGCCGC	AATGCCGTCT	GA	

BNSDOCID: &lt;WO 8957280A2 I &gt;

m136.pep

```

1  METNASILTA  TRLVFSAAAA  RTGIVPACFF  AFPADGLRFV  DDCLPVAVDI
51  RQCIRQLGFQ  FRQLAFCELO  TDSAVFLFV  NTAQCHDGIK  QLFKRFIIDG
101 FKPIGRHNIQ  TVKISIAPCV  KIAAAVFVFI  QPQIGOFFIR  HRGGCFHRHC
151 QNQPFDFGTF  GGGKLRFVAQ  HFGQPVERCQ  FVRPAQQRH  KTLNLVATHR
201 VALFAFGIQQ  FAQPPFGCFG  KFSGIHHFPF  QMGFAPYYRR  NAV*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 136 shows 85.6% identity over a 209 aa overlap with a predicted ORF (ORF 136.ng) from *N. gonorrhoeae*:

m136/g136

```

                                10      20      30      40
m136.pep      METNASILTATRLVFSAAAARTGIVPACFFAFPADGLRFVDDCLPV
g136           MEIRFQTAFLRLVQMKTNASILTATRLVFPAARTGIVPAGFFFPADGLRFVDDRLPV
                                10      20      30      40      50      60
m136.pep      AVDIRQCIRQLGFQFRQLAFCELOTD SAVFLFVVNTAQCHDGIKQLFKRFIIDGFKPIGR
g136           AVDVCQRVRQFGRKFRQLAFGELQADNAVFLFVVNAAHCHHGKQLFKRFIIGGFKPIGR
                                50      60      70      80      90     100
m136.pep      HNIQTVKISIAPCVKIAAAVFVFIQ P QIGOFFIRHRGGCFHRHCQNPFDGTFGGGKLR
g136           HNVQTVKIGVAPSVKIAAALAVVVEPQIGQLFIRHRGGCFHRHCQNPFDGTFGGGKLR
                                110     120     130     140     150     160
m136.pep      FVAQHFGQPVERCQFVRPAQQRHKTNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIH
g136           FVAQHFGQPVERCQFVRPAQQRHKTNLVATHRVALFAFGIQX
                                170     180     190     200     210     220
m136.pep      HFPFQMGFAPYYRRNAVX
g136           HFPFQMGFAPYYRRNAVX
                                230     240

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 527>:

a136.seq

```

1  ATGGAACAA  ACGCTTCAAT  TCTTACCGCA  ACACGCCTTG  TATTTTCTGC
51  CGCTGCCGCA  CGGACAGGGA  TCGTTCCTGC  CTGTTTTTTC  GCCTTCCCTG
101 CGGACGGTTT  GCGGCTTGTT  GATGACCGCC  TGCCAGTAGC  GGTAGATATC
151 CGCCAATGCA  TAAGGCAACT  CGGATTCCAG  TTCCGCCAGC  TCGCCTTCTG
201 TGAATTGCAG  ACGGATAGTG  CCGTTGTCCT  CTTGTCGTA  AATACGCCCC
251 AATGCCATGA  TGGGATAAAA  CAACTCTTCA  AACGCTTCAT  CATCGACGGC
301 TTCAAACCAA  TCGGTCGGCA  CAATATCCAA  ACCGTAAAGA  TAAGCATTGC
351 ACCATGTGTA  AAAATCGCTG  CCGCCGTCTT  CGTTTTTATA  CAGCCACAAA
401 TCGGGCAGTT  TTTTATCCGA  CATCGCGGCG  GTTGTTCATA  TCGCCATTGC
451 CAAAACCAGC  CGTTCGATTT  CGGAACGTT  GCGGCGGTA  AATTGCGATT
501 CGTCGCCCAA  CACTTCGGGC  AGCCAGTCGA  GCGGTGTCAA  TTTGTCCGGC
551 CCGCTCAACA  GCGCCGTCAT  AAAACCTTGA  ACCTCGTCGC  AACGCATCGT
601 GTTGCCCTGT  TCGCTTTTGG  CATCCAACAA  TTCGCTCAAC  CGCCGTTTGG
651 ATGCTTCGGT  AAATTTTCGG  GAATCCATCA  TTTTCCTTTT  CCAATGGGTT
701 TTGCGCCCTA  TTATAGTGGA  TTAAATTTAA  ATCAGGACAA  GGCGACGAAG
751 CCGCAGACAG  TACAAATAGT  ACGGCAAGGC  GAGGCAACGC  CGTACTGGTT
801 TAAATTTAAT  CCACTATATC  GCCGCAATGC  CGTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 528; ORF 136.a>:

a136.pep



```

1  METNASILTA  TRLVFSAAAA  RTGIVPACFF  AFPADGLRLV  DDRLPVAVDI
51  RQCIRQLGFQ  FRQLAFCELO  TDSAVVLFVV  NTAQCHDGIK  QLFKRFIIDG
101 FKPIGRHNIQ  TVKISIAPCV  KIAAAVFVFI  QPQIGQFFIR  HRGGCFHRHC
151 QNQPFDFGTF  GGGKLRVFAQ  HFGQPVERCQ  FVRPAQQRH  KTLNLVATHR
201 VALFAFGIQQ  FAQPPFGCFG  KFSGIHHFPF  PMGFAPYYSG  LNLNQDKATK
251 PQTQIVRQG  EATPYWFKFN  PLYRRNAV*

```

m136/a136 98.3% identity in 238 aa overlap

```

              10      20      30      40      50      60
m136.pep    METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRFVDDCLPVAVDIRQCIRQLGFQ
a136        METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRLVDDRLPVAVDIRQCIRQLGFQ
              10      20      30      40      50      60

              70      80      90     100     110     120
m136.pep    FRQLAFCELOTD SAVVLFVVNTAQCHDGIKQLFKRFIDGFKPIGRHNIQTVKISIAPCV
a136        FRQLAFCELOTD SAVVLFVVNTAQCHDGIKQLFKRFIDGFKPIGRHNIQTVKISIAPCV
              70      80      90     100     110     120

              130     140     150     160     170     180
m136.pep    KIAAAVFVFIQ P QIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ
a136        KIAAAVFVFIQ P QIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ
              130     140     150     160     170     180

              190     200     210     220     230     240
m136.pep    FVRPAQQRHKT LNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPFQMGFAPYYRR
a136        FVRPAQQRHKT LNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPFPMGFAPYYSG
              190     200     210     220     230     240

m136.pep    NAVX
a136        LNLNQDKATKPQTQIVRQGEATPYWFKFNPLYRRNAVX
              250     260     270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 529>:

```

g137.seq
1  ATGATTATCC ATCACcaATT CGATCCCGTC CTCATCAGTA TCGGCCCGCT
51  TGCCGTCCGC TGGTATGCCT TAAGCTACAT CCTCGGATTT ATTCTTTTTA
101 CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAAA
151 GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TGATTTTGGG
201 CGGACGCTTG GGCTATGTCC TGTTTTACAA ATTCTCCGAC TACCTCGCCC
251 ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301 GGCTTTTTGG GTGTAGTTAT TGCCATATGG TTGTTACGCC GCAAGCACGG
351 CATCGGCTTC CTCAAAGTGA TGGACACGGT CGCGCCGCTC GTTCCGCTGG
401 GTCTCGCTTC GGGACGTATC GGCAACTTTA TCAACGGCGA ACTTTGGGGA
451 CGCATTACCG ACATTAACGC ATTTTGGGCA ATGGGCTTCC CGCAAGCGCA
501 TTACGAAGAT GCCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAATGGC
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCCCTTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TTTCCAAAAA
651 ACCGCGCCCG ACCGGGCAGA CTGCGCGGCT TTTTCTCGGC GGCTACGGCG
701 TGTTCCGCTT TATTGCCGAA TTTGCGCGCC AACCCGACGA CTATCTCGGG
751 CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCAGATGAT
801 TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AAACAGCACT
851 GA

```

This corresponds to the amino acid sequence <SEQ ID 530; ORF 137.ng>:

g137.pep

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 531>:

m137.seq

1	ATGATTACCC	ATCCCCAATT	CGATCCCCTC	CTTATCAGTA	TCGGCCCGCT
51	TGCCGTCCGC	TGGTATGCCC	TAAGCTACAT	CCTCGGAATT	ATTCCTTTTA
101	CCTTCTCGG	CAGAAGGCGC	ATCGCGCAAG	GCTTGTCGGT	TTTTACCAAA
151	GAATCGCTTG	ACGACTTCCT	GACATGGGGC	ATTTTGGGCG	TAATTTTGGG
201	CGGGCGTTTG	GGTATCGTCC	TGTTTTACAA	GTTTTCCGAC	TACCTCGCCC
251	ATCCGCTTGA	TATTTTCAAG	GTATGGGAAG	GCGGAATGTC	GTCCACGGC
301	GGCTTTTTTG	GTGTAGTTAT	TGCCATACGG	TTGTTCCGGC	GCAAACACGG
351	CATCGGCTTC	CTCAAACCTGA	TGGATACGGT	CGCACCGCTC	GTTCCGCTGG
401	GTCTCGCTTC	GGGACGTATC	GGCAACTTCA	TCAACGGCGA	ATTTTGGGGA
451	CGCGTTTACCG	ACATCAACGC	ATTTTGGGCA	ATGGGCTTCC	CGCAGGCGCG
501	TTACGAAGAT	GCCGAAGCCG	CCGCGCACAA	TCGCTTTTGG	GCAGAATGGC
551	TGCAACAATA	CGGTATGCTG	CCGCGCTCATC	CCTCGCAGCT	TTATCAGTTT
601	GCACTTGAAG	GCATCTGCCT	GTTCAAGCTC	ATTTGGCTGT	TCTCTAAAAA
651	ACAGCGGTTCG	ACCGGACAAG	TCGCCTCGCT	CTTCCTCGGC	GCTCATGGCA
701	TATTCCGCTT	CATTGCCGAA	TTCGCACGCC	AACCCGACGA	CTATCTCGGG
751	CTGCTGACCT	TGGGCTGTGC	GATGGGGCAA	TGGTTGAGCG	TCCCGATGAT
801	TGTTTTGGGT	ATCGTCGGCT	TTGTCCGGTT	CGGCATGAAA	AAACAGCACT
851	GA				

m137.pgp

1	MITHPQFDPV	LISIGPLAVR	WYALS <del>Y</del> ILGF	ILFTFLGRRR	IAQGLSVFTK
51	ESLDDFLTWG	ILGVILGGRL	GYVLFYKFS	YLAHPLDIFK	VWEGGMSFHG
101	<u>GFLGVVIAIR</u>	<u>LFGRKHGIGF</u>	LKLM <del>DT</del> VAPL	VPLGLASGRI	GNFINGELWG
151	RVTDINAFWA	MGFPQARYED	AEAAAHNPLV	AEWLQQYGML	PRHPSQLYQF
201	<u>ALEGICLFTV</u>	<u>IWLFSKKQRS</u>	TQGVASFLFG	GYGIFRFAIE	FARQPD <del>Y</del> DLG
251	LLTLGLSMGQ	WLSVPMIVLG	IVGFRVFGMK	KOH*	

### Homology with a predicted ORF from *N. gonorrhoeae*

ORF 137 shows 95.4% identity over a 283 aa overlap with a predicted ORF (ORF 137.ng) from *N. gonorrhoeae*:

m137/g137

[illegible]

389

```

|||||
g137      AEWLQQYGMLPRHPSQLYQFALEGICLFAVVWLFSSKKPRPTGQTAALFLGGYGVRFRFAE
          190      200      210      220      230      240

          250      260      270      280
m137.pep  FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGVFVRFGMKKQHX
          |||||
g137      FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGVFVRFGMKKQHX
          250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 533>:

```

a137.seq
1  ATGATTACCC ATCCCAATT CGACCCCGTC CTTATCAGTA TCGGCCCGCT
51  TGGCGTCCGC TGGTATGCCC TAAGCTACAT CCTCGGATTT ATTCTTTTTA
101 CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAA
151 GAATCGCTCG ACGACTTCCT GACATGGGCG ATTTTGGGCG TAATTTTGGG
201 CGGGCGTTTG GGTACGTCC TGTTTTACAA GTTTTCCGAC TACCTCGCCC
251 ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301 GGCTTTTTTG GTGTAGTTAT TGCCATATGG TTGTTCCGTC GCAAACACGG
351 CATCGGCTTC CTAAACTGA TGGACACGGT CGCACCGCTC GTTCCACTGG
401 GTCTCGCTTC GGGACGTATC GGCAACTTCA TCAACGGCGA ACTTTGGGGA
451 CGCGTTACCG ACATCAACGC ATTTTGGGCA ATGGGCTTCC CGCAGGCGCG
501 TTACGAAGAC CTCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAATGGC
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCACTTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TCTCTAAAAA
651 ACAGCGGCCG ACCGACAAAG TCGCCTCACT CTTCTCGGCG GGCTACGGCA
701 TATTCCGCTT CATTGCCGAA TTTGCACGCC AACCCGACGA CTATCTCGGG
751 CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCAGATGAT
801 TGTTTTGGGT ATCGTCCGCT TTGTCCGGTT CGGCATGAAA AAACAGCACT
851 GA

```

This corresponds to the amino acid sequence <SEQ ID 534; ORF 137.a>:

```

a137.pep
1  MITHPQFDPV LISIGPLAVR WYALSYILGF ILFTFLGRRR IAQGLSVFTK
51  ESLDDFLTWG ILGVILGGRL GYVLFYKFS YLAHPLDIF VWEGGMSFHG
101 GFLGVVIAIW LFGRKHGIG LKLMDTVAPL VPLGLASGRI GNFINDELWG
151 RVTDINAFWA MGFPOARYED LEAAAHNPLW AEWLQQYGML PRHPSQLYQF
201 ALEGICLFAV VWLFSSKKQRP TGQVASLFLG GYGIFRFAE FARQPDDYL
251 LLTLGLSMGQ WLSVPMIVLG IVGFVRFGMK KQH*

```

m137/a137 98.2% identity in 283 aa overlap

```

          10      20      30      40      50      60
m137.pep  MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG
          |||||
a137      MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG
          10      20      30      40      50      60

          70      80      90      100     110     120
m137.pep  ILGVILGGRLGYVLFYKFSYLAHPLDIFKVWEGGMSFHGGFLGVVIAIRLFGGRKHGIGF
          |||||
a137      ILGVILGGRLGYVLFYKFSYLAHPLDIFKVWEGGMSFHGGFLGVVIAIRLFGGRKHGIGF
          70      80      90      100     110     120

          130     140     150     160     170     180
m137.pep  LKLMDTVAPLVPLGLASGRIGNFINDELWGGRVTDINAFWAMGFPOARYEDAEAAAHNPLW
          |||||
a137      LKLMDTVAPLVPLGLASGRIGNFINDELWGGRVTDINAFWAMGFPOARYEDLEAAAHNPLW
          130     140     150     160     170     180

          190     200     210     220     230     240
m137.pep  AEWLQQYGMLPRHPSQLYQFALEGICLFTVIWLFSSKKQRPSTGQVASLFLGGYGIFRFAE
          |||||
a137      AEWLQQYGMLPRHPSQLYQFALEGICLFAVVWLFSSKKQRPSTGQVASLFLGGYGIFRFAE

```

	190	200	210	220	230	240
m137.pep	250	260	270	280		
	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX					
a137	250	260	270	280		
	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 535>:

g138.seq

```

1  ATGGAGTTTG AAAACATTAT TTCCGCCGCGc gaCAAGGCGC GTATCCTTGC
51  CGAAGCACTG CCTTACAtcc gccgGTTTTTC CGGTTCCGTC GCCGTCATCA
101 AGTATGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
251 GCGAATTGT CCAAGGAATG CGCGTTACCG ACAAAGAGAC GATGGATATT
301 GTCGAAATGG TATTGGGCGG GCACGTCAAC AAGGAAATCG TGTGATGAT
351 TAACACATAT GGAGGGCACG CGGTCGGCGT GAGCGGGCGC GACGACCATT
401 TCATTAAGGC GAAGAACTT TTGGTCGATA CGCCCGAACA GAATAGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGCGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT GGCAGGCAAA
601 TTGGCGGAAG AATTGAACGC CGAAAACTC TTGATGATGA CGAAtatcgc
651 cgGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACCT acgCCGAAAC
701 GGATTGATGG GCTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCcgccgtc aACGGTGTGA AAGCCACGCA
801 CATCATCGAC GGCAGGTTGC CCAACGCGCT TTGCTGGAA ATCTTTACCG
851 ATGCCGGTAT CGGTCGATG ATTTTAGGCA GAGGGGAAGA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 536; ORF 138.ng>:

g138.pep

```

1  MEFENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
51  RDVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKETMDI
101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LVDTPEQNSV
151 DIGQVGTVES IDTGLVKGLI ERGCI PVVAP VGVGEKGEAF NINADLVAGK
201 LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDGLIA DGTLYGGMLP
251 KIASAVEAAV NGVKATHIID GRLPNALLLE IFTDAGIGSM ILGRGEDA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 537>:

m138.seq

```

1  ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
51  CGAAGCGCTG CCTTACATCC GCCGGTTTTTC CGGTTCCGTC GCCGTCATCA
101 AATACGCGCG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
251 GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGGC GATGGATATT
301 GTCGAAATGG TGTGGGCGG GCATGTCAAT AAAGAAATCG TGTGATGAT
351 TAACACATAT GGCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
401 TCATTAAGGC GAAGAACTT TTGATCGATA CGCCCGAACA GAATGGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGTGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT AGCAGGCAAA
601 TTGGCGGAAG AATTGAACGC CGAAAACTC TTGATGATGA CGAATATCGC
651 CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACCT ACGCCGAAAC
701 GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCCGCGTC AACGGTGTGA AAGCCACGCA
801 TATCATCGAC GGCAGGTTGC CCAACGCGCT TTGCTGGAA ATCTTTACCG
851 ATGCCGGTAT CGGTCGATG ATTTTGGGCG GTGGGAAGA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 538; ORF 138>:

m138.pep

```

1  MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA

```

### Homology with a predicted ORF from *N. gonorrhoeae*

m138/g138

	10	20	30	40	50	60
m138.pep	MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
g138	MEFENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m138.pep	IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKEAMDIVEMVLGGHVNKEIVSMINTY					
g138	IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKETMDIVEMVLGGHVNKEIVSMINTY					
	70	80	90	100	110	120
	130	140	150	160	170	180
m138.pep	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
g138	GGHAVGVSGRDDHFIKAKKLLVDTPEQNSVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m138.pep	VGVGEKGAEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTNLLTKLTPKRIDELIA					
g138	VGVGEKGAEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTNLLTKLTPKRIDGLIA					
	190	200	210	220	230	240
	250	260	270	280	290	299
m138.pep	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGGGEDAX					
g138	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGRGEDAX					
	250	260	270	280	290	

a138.seq

BNSDOCID: <WO\_8957280A2 | >

This corresponds to the amino acid sequence <SEQ ID 540; ORF 138.a>:

```
a138.pep
1  MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
51  RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKAMDI
101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
151 DIGQVGTVES IDTGLVKGLI ERGCI PVVAP VGVGEKGEAF NINADLVAGK
201 LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
251 KIASAVEAAV NGVKATHIID GRVNPALLLE IFTDAGIGSM ILGGGEDA*
```

m138/a138 99.7% identity in 298 aa overlap

```

10      20      30      40      50      60
m138.pep  MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG
a138      |||||
10      20      30      40      50      60
MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG

70      80      90      100     110     120
m138.pep  IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKAMDIVEMVLGGHVNKEIVSMINTY
a138      |||||
70      80      90      100     110     120
IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKAMDIVEMVLGGHVNKEIVSMINTY

130     140     150     160     170     180
m138.pep  GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCI PVVAP
a138      |||||
130     140     150     160     170     180
GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCI PVVAP

190     200     210     220     230     240
m138.pep  VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA
a138      |||||
190     200     210     220     230     240
VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA

250     260     270     280     290     299
m138.pep  DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGGGEDAX
a138      |||||
250     260     270     280     290
DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRVNPALLLEIFTDAGIGSMILGGGEDAX

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 541>:

```
g139.seq
1  ATGCGAACCA CCTCAACCTT CCCTACAAAA ACTTTCAAAC CGGCTGCCAT
51  GGCGTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAggc ggcggcggag
101 gcGGCACTTC TGCTCCCGAC TTTAATGCAG GCGGCACCGG TATCGGCAGC
151 AACAGCAGGG CAACGATAGC GGAATCAGCA GCAGTATCTT ACGCCGGTAT
201 AAAAAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCTGGGATG
251 ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAAAGCCCC CCGAATCTGC
301 ATACCGGAGA CTTTCAAAC CCAAATGACC AATATTAAGA ATATGATCAA
351 CCTCAAACCT GCAATTGAAG CAGGCTATAC AGGACGCGGG GTAGAGGTAG
401 GTATCGTCTGA TACAGGCGAA TCCGTCGGCA GCATATCCTT TCCCGAACTG
451 TATGGCAGAA AAGAACACGG CTATAACGAA AATTACAAAA ACAAAATTACA
501 AAAACTATAC GGCGTATATG CGGAAGGAAG CGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 542; ORF 138.ng>:

```
g139.pep
1  MRITSTFPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS
51  NSRATIAESA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKIKAPRIC
101 IPETFQTQMT NIKNMINLKP AIEAGYTGRG VEVGIVDTGE SVGSISFPPEL
151 YGRKEHGYNE NYKNKLQKLY GVYAEGSA*
```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 543>:

```
m139.seq
1  ATGCGAACGA CCCAACCTT CCCTACAAA ACTTTCAAAC GACTGCCAT
51  GGC GTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAGGC GGCGCGGAG
101 GCGGCACTTC TGCGCCCGAC TTCAATGCAG GCGGTACCGG TATCGGCAGC
151 AACAGCAGAG CAACAACAGC GAAATCAGCA GCGAGTACTT AGCCCGGTAT
201 CAAGAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCCGGATG
251 ACGTTGCGGT TACAGACAGG GATGCCAAA TCAATGCCCC CCCCCGAATC
301 TGCATACCGG AGACTTTCCA AACCCAAATG ACGCATtACA AGAATTGAT
351 CAACCTCAA CCTGCAATTG AAGCAGGCTA TACAGGACGC GGGGTAGAGG
401 TAGGTATCGT CGACACAGGC GAATCCGTAC GCAGCATATC CTTTCCCGAA
451 CTGTATGGCA GAAAGAACA CGGCTATAAC GAAAATTACG AAAAATAATA
501 CGGCGTATAT GCGGAAGGAA GCGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 544; ORF 138>:

```
m139.pep
1  MRTTPTFPTK  TFKPTAMALA  VATTLSACLG  GGGGGTSAPD  FNAGGTGIGS
51  NSRATTAKSA  AVSYAGIKNE  MCKDRSMLCA  GRDDVAVTDR  DAKINAPPRI
101 CIPETFQTQM  THYKNLINLK  PAIEAGYTGR  GVEVGIVDTG  ESVGSISFPE
151 LYGRKEHGYN  ENYEKLYGVY  AEGSA*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 138 shows 92.2% identity over a 179 aa overlap with a predicted ORF (ORF 138.ng) from *N. gonorrhoeae*:

```

m139/g139

                10          20          30          40          50          60
m139.pep  MRTTPTFPKTFKPTAMALAVATTLACLGGGGGGTSAPDFNAGGTGIGSNSRATTAKSA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g139      MTTTSTFPKTFKPAAMALAVATTLACLGGGGGGTSAPDFNAGGTGIGSNSRATIAESA
                10          20          30          40          50          60

                70          80          90          100         110         120
m139.pep  AVSYAGIKNEMCKDRSMLCAGRDDVAVTDTRDAKINAPPRICIPETFQTOMTHYKNLINLK
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g139      AVSYAGIKNEMCKDRSMLCAGRDDVAVTDTRDAKIKAP-RIICIPETFQTOMTNIKNMINLK
                70          80          90          100         110

                130         140         150         160         170
m139.pep  PAIEAGYTGRGVEVGIVDTGESVGSISFPPELYGRKEHGYNENY----EKLYGVYAEGSAX
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g139      PAIEAGYTGRGVEVGIVDTGESVGSISFPPELYGRKEHGYNENYKNKLQKLYGVYAEGSAX
          120          130          140          150          160          170

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 545>:

```

a139.seq
1  ATGCGAACGA CCCAACCTT CCCTACAAA ACTTTCAAAC CGGCTGCCAT
51  GGCCTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAGGC GCGGCGGAG
101 GCGGCACTTC TGCGCCCGAC TTCAATGCAG GCGGCACCGG TATCGGCAGC
151 AACAGCAGGG CAACAACAGC GAAATCAGCA GCAATATCTT AGCCCGGTAT
201 CAAGAACGAA ATGTGCAAAAG ACAGAAGCAT GCTCTGTGCC GCTCGGATG
251 ACGTTGCGGT TACAGACAGG GATGCCAAA TCAATGCCCC CCCCCGAATC
301 TGCATACCGG AGACTTTACA AACCCAAATG ACGCAT.ACA AGAATTTGAT
351 CAACCTCAAA CCTGCAATTG AAGCAGGCTA TACAGGACGC GGGGTAGAGG
401 TAGGTATCGT CGACACAGGC GAATCCGTAC GCAGCATATC CTTTCCCGAA
451 CTGTATGGCA GAAAAGAACA CGGCTATAAC GAAAATTAC. AAAAATAATA
501 CGGCGTATAT GCGGAAGGAA GCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 546; ORF 139.a>:

a139.pep  
1 MRTTPTEPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS

	10	20	30	40	50	60
m139.pep	MRTTPTFPKTKFKPTAMALAVAT	TL	SACLG	GGGGG	TSAPDFNAGGTGIGSNSRATTAKSA	
a139	:					
	10	20	30	40	50	60
	MRTTPTFPKTKFKPAAMALAVAT	TL	SACLG	GGGGG	TSAPDFNAGGTGIGSNSRATTAKSA	
	70	80	90	100	110	120
m139.pep	AVSYAGIKNEMCKDRSMLCAGR	DDVAVTDRDAKINAPPRICIPET	FTQTOMTHYKNLINLK			
a139	!:					
	70	80	90	100	110	120
	AISYAGIKNEMCKDRSMLCAGR	DDVAVTDRDAKINAPPRICIPET	LQTOMTHXKNLINLK			
	130	140	150	160	170	
m139.pep	PAIEAGYTGRGVEVGIVDTGES	VGSI	SFP	ELYGRKEHGYNENYKLYGVYAEGSAX		
a139						
	130	140	150	160	170	
	PAIEAGYTGRGVEVGIVDTGES	VGSI	SFP	ELYGRKEHGYNENYKLYGVYAEGSAX		

g140.seq

1	Atgtcgggcac	gCGGCAAGGG	GGCAGgctat	ctcAACAGTA	CCGGACGACa
51	TGTTCCCTTC	CTGAGTGCCG	CCAAAATCGG	GCAGGATTAT	TCTTTCTTCA
101	AAAATATCAA	AACCGACGGC	GGTCTGCTGG	CTTCCCTCGA	CAGCGTCGAA
151	AAAACAGCGG	GCAGTGAAGG	CGACACGCCG	TCCTATTATG	TCCGTGCGGG
201	CAATGCGCGA	CGGACTGCTT	CGGCAGCGGC	ACATTCCCGC	CCCGCCGGTC
251	TGAAACACGC	CGTAGAACAG	GGCGGCAGCA	ATCTGGAAAA	CTGATGGTCT
301	GAGCTGGATG	CCTCCGAATC	ATCCGCAACA	CCCGAGACGG	TTGAAACTGC
351	GGTCGCCGAC	CGCACAGATA	TGCCGGGCAT	CCGCCTACGG	CGCACAACCT
401	TCCGCACAGC	GGCAGCCGTA	CAGCATGCGA	ATACCGCCGA	CGGCGTACGC
451	aTCTTcaaCA	GTCTCGCCGC	TAccgTCTAT	GccgACAGTG	CCGCCGCCCA
501	TGccgATATG	CAGGACGCCC	GCCTGAAAGC	CGTATCGGAC	GGGTTGGTACC
551	ACAACGGTAC	GGGTCTGCGC	GTCAATCGCGC	AAACCCAAAC	GGACCGGTGGA
601	ACGTGGGAAC	AGGGCGGTGT	CGAAGGCAAA	ATGCGCGGCA	GTACCCAAAC
651	TATCGGCATT	GCCGCGAAAA	CCGCGCAAAA	TACGACAGCA	GCCGCCACAC
701	TGGGCATAGG	ACGCAGCACA	TGGAGCGAAA	ACAGTGCAAA	TGCAAAAACC
751	GACAGCATTa	GCTGTGTTGC	AGGCATACGG	CACGATGTGG	GCGATATCCG
801	CTATCTCAAA	GGCTGTGTTCT	CctaCGGACG	CTACAAAAAC	AGCATCAGCC
851	GCAGCACCGG	TGCGGATGAA	TATGCGGAAG	CGACGCTCAA	CGGCACGCTG
901	ATGCAGCTGG	GCGCACTGGG	TGGTGTCAAC	GTTCCGTTTTG	CCGCAACGGG
951	AGATTTGACG	GTTGAAGGCG	GTCTGCGCCA	CGACCTGCTC	AAACAGGATG
1001	CATTGCGCGA	AAAAGGcCagt	GCTTTGGGGCT	GGAGCGGCAA	CAGCCTCACT
1051	GAAGGCACAC	TAGTCCGACT	CGCGGGTCTG	AAACTGTCCG	AACCTTGTAG
1101	CGATAAAGCC	GTCCTGTCTG	CGACGGCGTG	CGTGGAAACG	GACCTGAACG
1151	GACGCGACTA	CGCGGTAACG	GGCGGCTTTA	CCGGCGCGGC	TGCAGCAACC
1201	GGCAAGACGG	GTGCACGCAA	TATGCCGCAC	ACCCGCCGGG	TTGCCGGTCT
1251	GGGGGTGGAT	GTCCAATTCTG	GCAACGGCTG	GAACGGCTTG	GCACGTTACA
1301	GCTACACCGG	TTCCAACACG	TACGGCAACC	ACAGCGGACA	AATCGCGCTA
1351	GGCTACCGGT	TCTGA			

g140.pep

1	MSARGKGAGY	LNSTGRHVPF	LSAAKIGQDY	SFFKNIKTGD	GLLASLDSVE
51	KTAGSEGDTF	SYVVRGNAA	RTASAAAHSA	PAGLKHAVEQ	GGSNLENLMV
101	ELDASESSAT	PETVETAVAD	RTDMPGIRLR	RTTFRATAAV	QHANTADGVR
151	IFNSLAATVY	ADSAAAHADM	QGRRLKAVSD	GLDHNGTGLR	VIAQTQQDGG
201	TWEQGGVEGK	MRGSTQTIGI	AAKTGENTTA	AATLGIGRST	WSENSANAKT



251 DSISLFAGIR HDVDIGYLYK GLFSYGRYKN SISRSTGADE YAEGSVNGTL  
 301 MQLGALGGVN VPFAATGDLT VEGGLRHDLL KQDAFAEKGS ALGWSGNSLT  
 351 EGTLVGLAGL KLSQPLSDKA VLSATAGVER DLNGRDYAVT GGFTGAAAT  
 401 GKTGARNMPH TRRVAGLGVD VEFNGWNGL ARYSYTGSKQ YGNHSGQIGV  
 451 GYRF\*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 549>:

m140.seq  
 1 ATGTCGGCAC GCGGCAAGGG GGCAGGCTAT CTCAACAGTA CCGGACGACG  
 51 TGTTCCTTC CTGAGTGCCG CCAAAATCGG GCAGGATTAT TCTTCTTCA  
 101 CAAACATCGA AACCGACGGC GGCCTGCTGG CTTCCTCGA CAGCGTCGAA  
 151 AAAACAGCGG GCAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGCGG  
 201 CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC  
 251 TGAAACACGC CGTAGAACAG GGCAGCAGCA ATCTGGAATA CCTGATGGTC  
 301 GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAACTGC  
 351 GGCAGCCGAC CGCACAGATA TGCCGGGCAT CCGCCCTAC GGCAGCAACTT  
 401 TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCCGA CGGTGTACGC  
 451 ATCTTCAACA GTCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA  
 501 TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC  
 551 ACAACGGCAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGGA  
 601 ACGTGGGAAC AGGGCGGTGT TGAAGGCAAA ATGCGCGGCA GTACCCAAAC  
 651 CGTCGGCATT GCCGCGAAAA CCGCGAAAA TACGACAGCA GCCGCCACAC  
 701 TGGGCATGGG ACGCAGCACA TGAGCGAAA ACAGTGCAA TGCAAAAACC  
 751 GACAGCATTG GTCTGTTTGC AGGCATACGG CACGATGCGG GCGATATCGG  
 801 CTATCTCAA GGCCTGTTCT CCTACGGACG CTACAAAAAC AGCATCAGCC  
 851 GCAGCACCGG TGCGGACGAA CATGCGGAAG GCAGCGTCAA CGGCACGCTG  
 901 ATGCAGCTGG GCGCACTGGG CGGTGTCAAC GTTCCGTTTG CCGCAACGGG  
 951 AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AACAGGATG  
 1001 CATTGCGCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCCTCACT  
 1051 GAAGGCACGC TGGTCGGACT CGCGGGTCTG AAGCTGTGCG AACCTTGAG  
 1101 CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG  
 1151 GACGCGACTA CACGGTAACG GCGGCTTTA CCGGCGCGAC TGCAGCAACC  
 1201 GGCAAGACGG GGGCAGCAA TATGCCGCAC ACCCGTCTGG TTGCCGGCCT  
 1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA  
 1301 GCTACGCCGG TTCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA  
 1351 GGCTACCGGT TCTGA

This corresponds to the amino acid sequence <SEQ ID 550; ORF 140>:

m140.pep  
 1 MSARGKGAGY LNSTGRRVPF LSAKIGQDY SFFTNIETDG GLLASLDSVE  
 51 KTAGSEGDITL SYVRRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV  
 101 ELDAESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR  
 151 IFNSLAATVY ADSTAAHADM QGRRLKAVSD GLDHNGTGLR VIAQTQDGG  
 201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT  
 251 DSISLFAGIR HDAGDIGYLYK GLFSYGRYKN SISRSTGADE HAEGSVNGTL  
 301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSLT  
 351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT  
 401 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV  
 451 GYRF\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 140 shows 94.5% identity over a 454 aa overlap with a predicted ORF (ORF 140.ng) from *N. gonorrhoeae*:

m140/g140

	10	20	30	40	50	60
m140.pep	MSARGKGAGYLNSTGRRVPFLSAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGDITL					
g140	MSARGKGAGYLNSTGRHVPFLSAKIGQDYSFFKNIKTDGGLLASLDSVEKTAGSEGDTP					
	10	20	30	40	50	60
	70	80	90	100	110	120

m140.pep SYVVRGNAARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAAD  
g140 SYVVRGNAARTASAAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAVAD  
70 80 90 100 110 120

m140.pep 130 140 150 160 170 180  
RTDMPGIRPYGATFRAAAVQHANAADGVRIFNLSLAATVYADSTAAHADMQGRRLLKAVSD  
g140 RTDMPGIRLRRTTFRATAAVQHANTADGVRIFNLSLAATVYADSAAAHADMQGRRLLKAVSD  
130 140 150 160 170 180

m140.pep 190 200 210 220 230 240  
GLDHNGTGLRVIAQTQDDGTTVEQGGVEGKMRGSTQTIGIAAKTGENTTAAATLGMGRST  
g140 GLDHNGTGLRVIAQTQDDGTTVEQGGVEGKMRGSTQTIGIAAKTGENTTAAATLGIGRST  
190 200 210 220 230 240

m140.pep 250 260 270 280 290 300  
WSENSANAKTDSISLFAIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL  
g140 WSENSANAKTDSISLFAIRHDVGDIGYLKGLFSYGRYKNSISRSTGADEYAEGSVNGTL  
250 260 270 280 290 300

m140.pep 310 320 330 340 350 360  
MQLGALGGVNVFPAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLTEGTLVGLAGL  
g140 MQLGALGGVNVFPAATGDLTVEGGLRHDLLKQDAFAEKGSALGWSGNSLTEGTLVGLAGL  
310 320 330 340 350 360

m140.pep 370 380 390 400 410 420  
KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD  
g140 KLSQPLSDKAVLSATAGVERDLNGRDYAVTGGFTGAAAAATGKTGARNMPHTRRVAGLGVD  
370 380 390 400 410 420

m140.pep 430 440 450  
VEFGNGWNLARYSYAGSKQYGNHSGRVGVGYRFX  
g140 VEFGNGWNLARYSYTGSKQYGNHSGQIGVGYRFX  
430 440 450

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 551>:

a140.seq  
1 ATGTCGGCAG GCGGTAAGGG GGCAGGCTAT CTCAACCGTA CCGGACAACG  
51 TGTTCCCTTC CTGAGTGCCG CCAAAATCGG GCGGGATTAT TCTTCTTCA  
101 CAAACATCGA AACCGACGGC GGTCTGCTGG CTCCCTCGA CAGCGTCGAA  
151 AAAACAGCGG GTAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGCGG  
201 CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC  
251 TGAACACGCG CGTAGAACAG GGCAGCAGCA ATCTGGAATA CCTGATGGTC  
301 GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC  
351 GGCCGCCGAC CGCACAGATA TGCCGGGCGAT CCGCCCTTAC GGCGCAACTT  
401 TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCCGA CCGGTGTACGC  
451 ATCTTCAACA ATCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA  
501 TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC  
551 ACAACGCTAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGGG  
601 ACGTGGGAAC AGGGCGGTGT TGAAGGCAAA ATGCGCGGCA GTACCCAAAC  
651 CGTCGGCATT GCCGCGAATA CCGGCGAATA TACGACAGCA GCCGCCACAC  
701 TGGGCATGGG ACACAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC  
751 GACAGCATTA GTCTGTTTGC AGGCATACGG CACGATGCGG GCGATATCGG  
801 CTATCTCAAA GGCCTGTTCT CCTACGGACG CTACAAAAC AGCATCAGCC  
851 GCAGCACCGG TGCGGACGAA CATGCGGAAG GCAGCGTCAA CCGCACGCTG  
901 ATGCAGCTGG GCGCACTGGG CCGTGTCAAC GTTCCGTTT CCGCAACGGG

```
951 AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AAACAGGATG
1001 CATTGCGCGA AAAAGGCACT GCTTTGGGCT GGAGCGGCAA CAGCATCACT
1051 GAAGGCACAC TGGTCGGACT CGCGGGTCTG AAGCTGTCGC AACCCTTGAG
1101 CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG
1151 GACGCGACTA CACGGTAACG GGCGGCTTTA CCGGCGCGAC TGCAGCAACC
1201 GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGCTGG TTGCCGCTCT
1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1301 GCTACGCCGG TTCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
1351 GGCTACCGGT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 552; ORF 140.a>:

```
a140.pep
1 MSAGGKGAGY LNRTGQRPVF LSAAKIGRDY SFFTNIETDG GLLASLDSVE
51 KTAGSEDTL SYVVRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV
101 ELDASESAT PETVETAAAD RTDMPGIRPY GATFRAAAAV QHANAADGVR
151 IFNNLAATVY ADSTAAHADM QGRRLLKAVSD GLDHNATGLR VIAQTQDGG
201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGHST WSENSANAKT
251 DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL QDAFAEKGS ALGWSGNSIT
351 EGTTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGTGATAAT
401 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV
451 GYRF*
```

m140/a140 98.2% identity in 454 aa overlap

```
m140.pep      10      20      30      40      50      60
MSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL
||| ||||| ||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a140          10      20      30      40      50      60
MSAGGKGAGYLNRTGQRPVFLLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL

m140.pep      70      80      90     100     110     120
SYVVRGNAAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDASESATPETVETAAAD
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a140          70      80      90     100     110     120
SYVVRGNAAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDASESATPETVETAAAD

m140.pep     130     140     150     160     170     180
RTDMPGIRPYGATFRAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLLKAVSD
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a140         130     140     150     160     170     180
RTDMPGIRPYGATFRAAAVQHANAADGVRIFNNLAATVYADSTAAHADMQGRRLLKAVSD

m140.pep     190     200     210     220     230     240
GLDHNATGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGRST
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a140         190     200     210     220     230     240
GLDHNATGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHST

m140.pep     250     260     270     280     290     300
WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a140         250     260     270     280     290     300
WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL

m140.pep     310     320     330     340     350     360
MQLGALGGVNVFPFAATGDLTVEGGLRYDLLQDAFAEKGSALGWSGNSLTEGTLVGLAGL
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a140         310     320     330     340     350     360
MQLGALGGVNVFPFAATGDLTVEGGLRYDLLQDAFAEKGSALGWSGNSITEGTLVGLAGL

m140.pep     370     380     390     400     410     420
KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a140         370     380     390     400     410     420
KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD
```

	370	380	390	400	410	420
	430	440	450			
m140.pep	VEFGNGWNLARYSYAGSKQYGNHSGRVGVGYRFX					
a140	VEFGNGWNLARYSYAGSKQYGNHSGRVGVGYRFX					
	430	440	450			

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 553>:

g141.seq

```

1  atgagcttca aaAccgATGC CGAAACCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTGGGTTT GAACGTTGAC AACATTGAGC
101 CTTACGGTCA TTACAAAGCC AAAATCAATC CTGCCGAAGC GTTCAAGCTG
151 CCGCAAAAAC AAGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
201 GCGGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
251 GCCATATCGG CAAAGACTCT GTGATTGCTT TGC GCGAGCC TTCTTTGGGT
301 CCGGTGTTCC GCGTGAAAGG CGGCGCGGCA GCGGCGGGCT ACGCGCAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGCGAC TTCCACGCCA
401 TCGGTGCGGC GAATAACCTC CTCGCGGCCA TGCTCGACAA CCATATCTAC
451 CAAGGTAACG AGTTGAACAT CGACCCCAA CGCGTGCTGT GCGGCGCGCT
501 GGTGCGATAT AACGACCGCC AGTTGCGCAA CATCATCGAC GGTATGGGCA
551 AGCCTGTtga cggCGTGATG CGtcccGACG GCTTCGACAT CACCGTCGCC
601 TCCGAAGTga tggcgGTATT CTGCCTTGCC AAAGACATCA GCGATTTGAA
651 AGAGCGTttt gGCAATATTC TCGTCGCCTA CGCCAAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTGAAG GCACACGGCG CGATGGCGGC ATTGCTAAAA
751 GATGCGATTA AGCCCAATTT GGTGCAAACC ATCGAAGGCA CTCGGCCTT
801 TGTACACGGC GGCCCGTTCC CCAACATCGC CCACGGCTGC AACTCCGTTA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
901 GGCTTCGGCG CGGACTTGGG TGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CCTTGCCGGT TTGAAACCTG ATGCGGCAGT CGTCGTGGCG ACTGTCCGCG
1001 CCCTGAAATA CAACGCGGCG GTGGAACGCG CCAACCTTGG TGAAGAAAC
1051 CTCGAAGCCT TGGCAAAAGG TTTGCCCAAC CTGTTGAAAC ACATTTCCAA
1101 CCTGAAAAAC GTATTGCGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
1151 TGTCCGACTC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GCGGCGCGGG
1251 CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA TGCCATCGAC AACCAACCTA
1301 ATAACCTCGG TTTGCGCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGAGTT CCCAAAAAGT GTACGGCGCG GAAGATGTCG ATTTTCAGCGC
1401 GGAAGCGTCT GCCGAAATCG CCTCGCTGGA AAAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATATT CATTGAGCGA CAACGCCAAA
1501 CTCTTGGGCT GCCCGAAGG CTTCCGCATC GCCGTACGCG GTATCACTGT
1551 TTCCGCGGCG GCGGGCTTCA TCGTTGCGTT GTGCGGCAAT ATGATGAAAA
1601 TGCCGGGCGT GCCGAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGAA
1651 CACGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 554; ORF 141.ng>:

g141.pep

```

1  MSFKTDAETA QSSTMRPIGE IAAKLGLNVD NIEPYGHYKA KINPAEAFKL
51  PQKQGRLLIV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
101 PVFVGKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANNL LAAMLDNHIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPDVGVM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERF GNILVAYAKD GSPVYAKDLK AHGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LEALAKGLPN LLKHISNLKN VFGLPVVVAL NRVSDSDAE LAMIEKACAE
401 HGVEVSLTEV WGKGAGGAD LARKVVNAID NQPNNFGFAY DVELGIKDKI
451 RAI AQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPGEGFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDE
551 HGVHGLF*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 555>:

m141.seq

```

1  ATGAGCTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTTGGTCT GAATGCCGAC AACATTGAGC
101 CTTACGGTCA TTACAAGGCG AAAATCAATC CTGCCGAAGC GTTCAAACCTG
151 CCGCAAAAC AGGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
201 GGCGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCGTTGC
251 GCCACATCGG CAAAGATGCC GTGATTGCCG TGCGCGAACC TTCTCTGGGG
301 CCGGTGTTTCG GCGTGAAAGG CGGCGCGGCA GCGGCGGGCT ATGCCCAAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTCACGCCA
401 TCGGTGCGGC AAATAATCTG CTTGCCGCGA TGCTCGACAA CCATATCTAC
451 CAAGGCAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GCGGCGCGCT
501 GGTCGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGTA
551 AACCCGTTGA CGGCGTGATG CGTCTGACG GTTTCGATAT TACCGTTGCT
601 TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTTGAA
651 AGAGCGTTTG GGCAACATCC TTGTGCGCTA CGCCAAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
751 TTAGACGCTT AGCCCAACTT GGTGCAAAAC ATCGAAGGCA CGCCCGCCTT
801 CGTACACGGC GGCCCGTTTCG CCAACATCGC CCACGCTGC AACTCCGTAA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCGT AACCGAAGCA
901 GGCTTCGGCG CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CCTTGCCGGT TTGAAACCTG ATGCGGCTGT TGTCGTGGCG ACTGTCCGCG
1001 CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAA
1051 TTAGACGCTT TGGAAAAAGG TTTGCCCAAC CTGCTGAAAC ACATTTCCAA
1101 CCTGAAAAAC GTATTCGGAC TGCCCGTCGT CGTTGCGCTC AACCGTTTCG
1151 TGTCCGACGC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
1251 CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA CGCCATTGAA AGTCAAACCA
1301 ATAACCTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCATGTT CCAAAAAAGT GTACGGCGCG GAAGATGTTG ATTTACGCGC
1401 GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
1501 CTGTTGGGCT GCGCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCGT
1551 TTCCGAGGCG GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601 TGCCCGGCCT GCGCAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGCA
1651 GAAGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 556; ORF 141>:

m141.pep

```

1  MSFKTDAEIA QSSTMRPIGE IAAKLGLNAD NIEPYGHYKA KINPAEAFKL
51  PQKQGRLLIV TAINPTPAGE GKTTVTIGLA DALRHIGKDA VIALREPSLG
101 PVFVGKGGAA GGGYAQVLPD EDINLHFTGD FHAIGAANL LAAMLNDHIY
151 QGNELNIDPK RVLWRRVDM NDRQLRNIID GMGKPDGVM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRVSDADAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVVNAIE SQTNNFGFAY DVELGIKDKI
451 RAIQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
551 EGVIHGLF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 141 shows 97.5% identity over a 558 aa overlap with a predicted ORF (ORF 141.ng) from *N. gonorrhoeae*:

m141/g141

```

          10          20          30          40          50          60
m141.pep  MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKL PQKQGRLLIV
          |||||  |||||  |||||  |||||  |||||  |||||
g141      MSFKTDAETAQSSTMRPIGEIAAKLGLNVNDNIEPYGHYKAKINPAEAFKL PQKQGRLLIV
          10          20          30          40          50          60

```

m141.pep	70	80	90	100	110	120
	TAINPTPAGEGKTTVTIGLADALRHIGKDAVIALREPSLGPVFGVKGAAGGGYAQVLPM					
g141	:					
	70	80	90	100	110	120
	TAINPTPAGEGKTTVTIGLADALRHIGKDSVIALREPSLGPVFGVKGAAGGGYAQVLPM					
	70	80	90	100	110	120
m141.pep	130	140	150	160	170	180
	EDINLHFTGDFHAIGAANNLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDMMNDRLRNIIID					
g141	:					
	130	140	150	160	170	180
	EDINLHFTGDFHAIGAANNLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDMMNDRLRNIIID					
	130	140	150	160	170	180
m141.pep	190	200	210	220	230	240
	GMGKPVVDGVMRPDGFDTIVASEVMAVFCCLAKDISDLKERLGNILVAYAKDGSFPVYAKDLK					
g141	:					
	190	200	210	220	230	240
	GMGKPVVDGVMRPDGFDTIVASEVMAVFCCLAKDISDLKERFGNILVAYAKDGSFPVYAKDLK					
	190	200	210	220	230	240
m141.pep	250	260	270	280	290	300
	ANGAMAALLKDAIKPNLVQTIIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
g141	:					
	250	260	270	280	290	300
	AHGAMAALLKDAIKPNLVQTIIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
	250	260	270	280	290	300
m141.pep	310	320	330	340	350	360
	GFGADLGAEKFCDIKCRLAGLKPDAVVVATVRALKYNGGVERANLGEENLDALEKGLPN					
g141	:					
	310	320	330	340	350	360
	GFGADLGAEKFCDIKCRLAGLKPDAVVVATVRALKYNGGVERANLGEENLEALAKGLPN					
	310	320	330	340	350	360
m141.pep	370	380	390	400	410	420
	LLKHISNLKNVFGPLPVVVALNRFVSDADAEELAMIEKACAEHGVESLTVWGKGGAGGAD					
g141	:					
	370	380	390	400	410	420
	LLKHISNLKNVFGPLPVVVALNRFVSDADAEELAMIEKACAEHGVESLTVWGKGGAGGAD					
	370	380	390	400	410	420
m141.pep	430	440	450	460	470	480
	LARKVVNAIESQTNNGFGAYDVELGIKDKIRAIQAQVYGAEDVDFSAEASAEIASLEKLG					
g141	:					
	430	440	450	460	470	480
	LARKVVNAIDNQPNNGFGAYDVELGIKDKIRAIQAQVYGAEDVDFSAEASAEIASLEKLG					
	430	440	450	460	470	480
m141.pep	490	500	510	520	530	540
	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
g141	:					
	490	500	510	520	530	540
	LDKMPICMAKTQYSLSDNAKLLGCPGEGFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
	490	500	510	520	530	540
m141.pep	550	559				
	PAAEKIDVDAEGVIHGLFX					
g141	:					
	PAAEKIDVDEHGVHGLFX					
	550					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 557>:

```

a141.seq
1  ATGAGTTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTGGGTTT GAACGTGAC AACATTGAGC
101 CTTACGGTCA TTACAAAGCC AAAATCAATC CTGCCGAAGC GTTCAAACCTG
151 CCGCAAAAAC AGGGCAGGCT GATTTGGTT ACCGCCATCA ACCCGACTCC

```

```

201  GCGCGGCGAA  GGTAAAACCA  CCGTAACCAT  CGGTTTGGCG  GACGCATTGC
251  GCCATATCGG  CAAAGACTCT  GTGATTGCTT  TGCGCGAGCC  TTCTTTGGGT
301  CCGGTGTTTC  GCGTGAAAGG  CGGCGCGGCA  GGCGGCGGCT  ATGCCCAAGT
351  TTTGCCGATG  GAAGACATCA  ACCTGCACTT  CACCGGAGAT  TTTCACGCCA
401  TCGGTGCGGC  AAATAATCTG  CTTGCCGCGA  TGCTCGACAA  CCATATCTAC
451  CAAGGCAACG  AGTTGAACAT  CGACCCCAAA  CGCGTGCTGT  GGC GGCGCGT
501  GGTGCGATAT  AACGACCGCC  AGTTGCGCAA  CATCATCGAC  GGCATGGGCA
551  AGCCTGTTGA  CGGCGTGATG  CGTCTGACG  GTTCGATAT  TACCGTTGCT
601  TCCGAAGTGA  TGGCGGTATT  CTGTCTTGCC  AAAGACATCA  GCGATTTGAA
651  AGAGCGTTTG  GGCAACATCC  TTGTCGCCTA  CGCCAAAGAC  GGCAGCCCCG
701  TTTACGCCAA  AGATTTGAAA  GCGAATGGCG  CGATGGCGGC  ATTGCTTAAA
751  GATGCGATTA  AGCCCAACTT  GGTGCAAACC  ATCGAAGGCA  CGCCCGCCTT
801  CGTACACGGC  GGCCCGTTTC  CCAACATCGC  CCACGGCTGC  AACTCCGTAA
851  CCGCAACCCG  TCTGGCGAAA  CACCTTGCCG  ATTACGCCGT  AACCGAAGCA
901  GGCTTCGGCG  CGGACTTGCG  CGCGGAAAAA  TTCTGCGACA  TCAAATGCCG
951  CCTTGCCGGT  TTGAAACCTG  ATGCGGCTGT  TGTGCTGGCG  ACTGTCCGCG
1001 CTTTGAATA  TAACGGCGGC  GTGGAACGCG  CCAACCTCGG  CGAAGAAAAA
1051 TTAGACGCTT  TGGAAAAAGG  TTTGCCCAAC  CTGCTGAAAC  ACATTTCCTA
1101 CCTGAAAAAC  GTATTTCGGC  TGCCCGTCGT  CGTTGCGCTC  AACCCTTCG
1151 TGTCCGACTC  CGATGCCGAG  TTGGCGATGA  TTGAAAAAGC  CTGTGCCGAA
1201 CACGGCGTTG  AAGTTTCCCT  GACCGAAGTG  TGGGGCAAAG  GTGGTGCGGG
1251 CGGCGCGGAT  TTGGCGCGCA  AAGTCGTCAA  CGCCATTGAA  AGTCAAACCA
1301 ATAATTCGG  TTTGCGCTAC  GATGTCGAGT  TGGGCATCAA  AGACAAAATC
1351 CTGCGATTG  CCCAAAAAGT  GTACGGCGCG  GAAGATGTTG  ATTTACGCGC
1401 GGAAGCGTCT  GCCGAAATCG  CTTCACTGGA  AAAACTGGGC  TTGGACAAAA
1451 TGCCGATCTG  CATGGCGAAA  ACCCAATACT  CTTTGAGCGA  CAACGCCAAA
1501 CTGTTGGGCT  GCGCCGAAGA  CTTCCGCATC  GCCGTGCGCG  GCATCACCGT
1551 TTCCGCAGG  GCAGGTTTCA  TCGTCGCCCT  GTGCGGCAAC  ATGATGAAAA
1601 TGCCCGGCT  GCCCAAAGTT  CCGGCTGCGG  AGAAAATCGA  TGTGGACGCA
1651 GAAGGCGTGA  TTCACGGCTT  GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 558; ORF 141.a>:

```

a141.pep
  1  MSFKTDAEIA  QSSTMRPIGE  IAAKLGLNVD  NIEPYGHYKA  KINPAEAFKL
 51  PQKQGRILIV  TAINPTPAGE  GKT TVTIGLA  DALRHIGKDS  VIALREPSLG
101  PVFVGKGGAA  GGGYAQVLP  EDINLHFTGD  FHAIGAANNL  LAAMLNDHIY
151  QGNELNIDPK  RVLWRRVDM  NDRQLRNIID  GMGKPV DGV  RPDGFEDITVA
201  SEVMAVFCLA  KDISDLKERL  GNILVAYAKD  GSPVYAKDLK  ANGAMAALLK
251  DAIKPNLVQT  IEGTPAFVHG  GPFANIAHGC  NSVTATRLAK  HLADYAVTEA
301  GFGADLGAEK  FCDIKCRLAG  LKPDAAVVVA  TVRALKYNGG  VERANLGEEN
351  LDALEKGLPN  LLKHISNLKN  VFGLPVVVAL  NRFVSDSDAE  LAMIEKACAE
401  HGVEVSLTEV  WGKGGAGGAD  LARKVVNAIE  SQTNNFGFAY  DVELGIKDKI
451  RAI AQKVYGA  EDVDFSAEAS  AEIASLEKLG  LDKMPICMAK  TOYSLSDNAK
501  LLGCPEDFRI  AVRGITVSAG  AGFIVALCGN  MMKMPGLPKV  PAAEKIDVDA
551  EGVIHGLF*

```

m141/a141 99.5% identity in 558 aa overlap

```

          10      20      30      40      50      60
m141.pep  MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKL PQKQGRILIV
          |||
a141       MSFKTDAEIAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKL PQKQGRILIV
          10      20      30      40      50      60

          70      80      90     100     110     120
m141.pep  TAINPTPAGEGKTTVTIGLADALRHIGKDAVIALREPSLGPVFGVKGGAGGGYAQVLP
          |||
a141       TAINPTPAGEGKTTVTIGLADALRHIGKDSVIALREPSLGPVFGVKGGAGGGYAQVLP
          70      80      90     100     110     120

          130     140     150     160     170     180
m141.pep  EDINLHFTGDFHAIGAANLLAAMLNDHIYQGNELNIDPKRVLWRRVDMNDRQLRNIID
          |||
a141       EDINLHFTGDFHAIGAANLLAAMLNDHIYQGNELNIDPKRVLWRRVDMNDRQLRNIID
          130     140     150     160     170     180

```

m141.pep	190	200	210	220	230	240
	GMGKPVGDGVMRPDGF DITVASEVMAVFCLAKDISDLKERLGNILVAYAKDGSPVYAKDLK					
a141	GMGKPVGDGVMRPDGF DITVASEVMAVFCLAKDISDLKERLGNILVAYAKDGSPVYAKDLK					
	190	200	210	220	230	240
m141.pep	250	260	270	280	290	300
	ANGAMAALLKDAIKPNLVQTI EGT PAFVHGGPFANIAHG CNSVTATRLAKHLADYAVTEA					
a141	ANGAMAALLKDAIKPNLVQTI EGT PAFVHGGPFANIAHG CNSVTATRLAKHLADYAVTEA					
	250	260	270	280	290	300
m141.pep	310	320	330	340	350	360
	GFGADLGA EKFC DIKRLAGLKPDA AVV VATVRALKYNGGVERANLGEENLDALEKGLPN					
a141	GFGADLGA EKFC DIKRLAGLKPDA AVV VATVRALKYNGGVERANLGEENLDALEKGLPN					
	310	320	330	340	350	360
m141.pep	370	380	390	400	410	420
	LLKHISNLKNVFGLPVVVALNRFVSDADAELAMIEKACAEHGVEVSLTEVWGKGAGGAD					
a141	LLKHISNLKNVFGLPVVVALNRFVSDADAELAMIEKACAEHGVEVSLTEVWGKGAGGAD					
	370	380	390	400	410	420
m141.pep	430	440	450	460	470	480
	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAI AQKVYGAEDVDFSAEASAEIASLEKLG					
a141	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAI AQKVYGAEDVDFSAEASAEIASLEKLG					
	430	440	450	460	470	480
m141.pep	490	500	510	520	530	540
	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
a141	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
	490	500	510	520	530	540
m141.pep	550	559				
	PAAEKIDVDAEGVIHGLFX					
a141	PAAEKIDVDAEGVIHGLFX					
	550					

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 559>:

```

g142.seq
1   ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCGTGC AGGTGCGCCA
51  ACGCGCCTTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAAATATGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTT
151 GGCAACATCC TGATGTTTCGT CCGCCAGCAT ATTGATGCAG AGGCTGCCGT
201 TTTCCGACAG GATcggaATG AttcgCGCAC TCCGGTTTAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCGGT AACCGGCGCA ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCAC
351 AGGATGCCAT CGCATCACGG AACGAAGTTT GAAAAGTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC CGTCCCTAT ATAAGAATGC TGCACACAAG
451 GCATCCCCC ATGTGCAGCA GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 560; ORF 142.ng>:

```

g142.pep
1   MRADFMFADN MPVQVRQRAF YFKLSRFAAM PNMVGKPLFG RQAGQPGKMF
51  GNILMFVRQH IDAEAAVFRQ DRNDSRTPVY AQHHGRRLLVG NRRNRHHCNA
101 VTPCRTVCRD DMNACRTGCH RITERSLKSF LQIRHFSPLN RPLYKNAAHK
151 ASPHVQOF*

```



The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 561>:

```
m142.seq
1  ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCCTGC AGGTGCGCCA
51  ACGCGCCCTC TATTTCAAGT TGTCCCCTTT TGCCGCGATG CCAGATGTGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAATGTTC
151 GGCAACATCC TGATGTTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCGGT AACCGGCGCG ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC
351 AAGATGCCAT CGCATCACGG AACGAAGTTT GAAAATTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG
451 GCATCCCCcC ATGTGCAGCA GTTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 562; ORF 142>:

```
m142.pep
1  MRADFMFADN MPVQVRQAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF
51  GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLVG NRRDRRHENA
101 VTPCRTVCRD DMNACRARCH RITERSLKIF LQIRHFSPLN CPLYKNAAHK
151 ASPHVQQF*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 142 shows 93.7% identity over a 158 aa overlap with a predicted ORF (ORF 142.ng) from *N. gonorrhoeae*:

```
m142/g142
10      20      30      40      50      60
m142.pep MRADFMFADNMPVQVRQALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMFNGNILMFVRQR
|||||
g142      MRADFMFADNMPVQVRQAFYFKLSRFAAMPNMVGKPLFGRQAGQPGKMFNGNILMFVRQH
10      20      30      40      50      60

70      80      90      100     110     120
m142.pep IDAEAAVFRQDRNDSRTPVDAQHHGRRLVGNRRDRRHENAVTPCRTVCRDDMNACRARCH
|||||
g142      IDAEAAVFRQDRNDSRTPVYAHGRRLVGNRRDRRHENAVTPCRTVCRDDMNACRTGCH
70      80      90      100     110     120

130     140     150     159
m142.pep RITERSLKIFLQIRHFSPLNCPYKNAAHKASPHVQQFX
|||||
g142      RITERSLKSFQIRHFSPLNRPYKNAAHKASPHVQQFX
130     140     150
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 563>:

```
a142.seq
1  ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCCTGC AGGTGCGCCA
51  ACGCGCCCTC TATTTCAAGT TGTCCCCTTT TGCCGCGATG CCAGATGTGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAATGTTC
151 GGCAACATCC TGATGTTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCGGT AACCGGCGCA ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC
351 AGGATGCCAT CGCATCACGG AACGAAGTTT GAAAAGTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG
451 GCACCCCCCA TGTGCAGCAG TTCTGATTCA AAAAGCCGTC GGTCCGACAT
501 TTCCGCGCGT TACGGCGTAT TACGAGTTCA ACGCATCCTC GATTTTGGCA
551 AGTTCTGCCA ACAGGTCTTT AAGCAGCAGC ATTTTCTCGC GGCCAGCAC
601 TTCCTCGATA CGGTCGTAAC GCTCGTCCAC TTCTTCGCCG ATTTCTCAT
651 ACAGCTTCTC GCCCTCGGCA GTCAGCTTCA GAAAAACACG TCGTTGGTCG
701 TTGGAAGGTT TCAGGCGGAC AACCAGACCC GCTTTTCAA GCGGGGTCAG
751 GATACCGGTC AGGCTGGGGC GCAAATGCA CGCTGATTG GCCAAATCTT
```

801 GAAAGTCCAG CGTGCCGTTT TCCGCCAAAA GACGGATAAT CCGCCATTGC  
851 TGATCGGTAA TATTCGCCTG ATTCAGAATA GGCCTGAATT GGGTCATCAG  
901 GGCTTCCCTT GCCTGTATCA GACCGATATT GATAGACGCA TGTTTTGA

This corresponds to the amino acid sequence <SEQ ID 564; ORF 142.a>:

a142.pep  
1 MRADFMFADN MPVQVRQAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF  
51 GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLLV NRRNRRHCNA  
101 VTPCRTVCRD DMNACRTGCH RITERSLKSF LQIRHFSPLN CPLYKNAAHK  
151 APPMCSSSDS KSRRSDISAR YGVLRVQRIL DFGKFCQQVF KQOHFLAAQH  
201 FLDSVVTLVH FFADFLIQLL ALGSQLOKNT SLVVGREFQAD NQTRFFKAGQ  
251 DTGQAGAQNA RLIRQILKVQ RAVFRQKTDN PPLLIGNIRL IQNRPELGHQ  
301 GFPCLYQTDI DRRMF\*

m142/a142 96.1% identity in 153 aa overlap

m142.pep	10	20	30	40	50	60
	MRADFMFADNMPVQVRQALYFKLSRFAAMPD	VVGKPLFGRQAGQPGKMF	GNILMFVRQR			
a142	MRADFMFADNMPVQVRQALYFKLSRFAAMPD	VVGKPLFGRQAGQPGKMF	GNILMFVRQR			
	10	20	30	40	50	60
m142.pep	70	80	90	100	110	120
	IDAEAAVFRQDRNDSRTPVDAQHHGRRLLVGNRRDRRHCNAVTPCRTVCRDDMNACRARCH					
a142	IDAEAAVFRQDRNDSRTPVDAQHHGRRLLVGNRRDRRHCNAVTPCRTVCRDDMNACRARCH					
	70	80	90	100	110	120
m142.pep	130	140	150	159		
	RITERSLKIFLQIRHFSPLNCPPLYKNAAHKASPHVQQFX					
a142	RITERSLKIFLQIRHFSPLNCPPLYKNAAHKASPHVQQFX					
	130	140	150	160	170	180
a142	DFGKFCQQVFQKQOHFLAAQHFLDSVVTLVHFFADFLIQLLALGSQLOKNTSLVVGREFQAD					
	190	200	210	220	230	240

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 565>:

g143.seq  
1 ATGTTGAGCT TCGGCTATCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAAG  
51 CTCGCAGATG AGCCGCATT TTTCAAACGCT AGGCGCAGAC CCGCACAAAT  
101 TGGGCTGGTT TTTCATCCTG CCGCCGCTGG CCGGGATGCT GGTTCAGCCG  
151 ATAGTGgGCT ACTACTCAGA CCGCACTTGG AAGCCGCGCT TGGGCGGCCG  
201 CCGCCTGCCG TATCTGCTTT ACGGCACGCT GATTGCGGTC ATCGTGATGA  
251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTGCGTGGCG  
301 GCCTTGTCGT TCGGCGCGCT GATGATTGCG CTGTTGGACG TGTCGTCGAA  
351 TATGGCGATG CAGCCGTTTA AGATGATGGT CCGCGATATG GTCAACGAGG  
401 AGCAGAAAAG CTACGCCTAC GGGATTCAAA GTTTCTTAGC GAATACGGAC  
451 GCGGTTGTGG CAGCGATTCT GCCGTTTGTG TTcgcgata TCGGTTTGGC  
501 GAACACTGCC GAGAAAGGCG TTGTGCCACA AACCCTGGTC GTAGCATTCT  
551 ATGTGGGTGC GCGGTTACTG ATTATTACCA GTGCGTTCAC AATCTCCAAA  
601 GTCAAAGAAT ACGACCCGGA AACCTACGCC CGTTACCACG GCATCGATGT  
651 CGCCGCGAAT CAGGAAAAAG CCAACTGGTT CGAACTCTTA AAAACCGCGC  
701 CTAAAGTGTT TTGGACGGTT ACTCCGGTAC AGTTTTTCTG CTGGTTCGCC  
751 TTCCGGTATA TGTGGACTTA CTCGGCAGGC GCGATTGCAG AAAACGTCTG  
801 GCACACTACC GATGCGTCTT CCGTAGGCCA TCAGGAGGCG GGCAACCGGT  
851 ACGGCGTTTT GCGGCGGGTG TAGTCGGTTG CCGCGGTGAT TTGTTCTGTTT  
901 ATTCTGGCAA AAGTACCGAA TAAATACCAT AAGGCGGGTT ATTTCTGGCTG  
951 TTTGGCTTTG GCGCGGCTCG GTTTCTTCTC TATCTTCTTC ATCTACAATC  
1001 AATACGCACT CATCCTGTCT TATATCTTAA TCGGCATCGC TTGGGCGGGC  
1051 ATTATCACTT ATCCGCTGAC GATTGTGGCC AACGCTTTGT CCGGCAACAA  
1101 CATGGATACT TATTTGGGCC TGTttaacgg ctctgtCTGT ATGCcgcaaa  
1151 tcgTcgctTC GctgttgAGT TTCGTGCTTT TCCCGATGCT GGGCGGCCAT

1201 CAGGCAACCA TGTTCTTGGT TGCAGGCGCA GTCTTGCTGC TGGGAGCCTT  
1251 CTCAGTCTGT CTGATTAAAG AGATCCACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 566; ORF 143.ng>:

g143.pep  
1 MLSFGYLGVO TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP  
51 IVGYSDRTW KPRLGGRRLP YLLYGTIAV IVMILMPNSG SFGFGYASLA  
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKSYAY GIQSFLANTD  
151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL IITSAFTISK  
201 VKEYDPETYA RYHGIDVAAN QEKANWFELL KTAPKVFMTV TPVQFFCWFA  
251 FRYMWTYSAG AIAENVWHTT DASSVGHQEA GNRYGVLAHV \*SVAAVICSF  
301 ILAKVPNKYH KAGYFGCLAL GALGFFSIFV IYNQYALILS YILIGIAWAG  
351 IITYPLTIVA NALSGKHMGT YLGLFNGSVC MPQIVASLLS FVLFPMLGGH  
401 QATMFLVAGA VLLLGAFSVC LIKEIHGGV\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 567>:

m143.seq  
1 ATGCTCAGTT TCGGCTTTCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAAG  
51 CTCGCAAATG AGCCGCATTT TTCAAACGCT AGGCGCAGAC CCGCACAATT  
101 TGGGCTGGTT TTTCATCCTG CCGCCGCTGG CGGGGATGCT GGTGCAGCCG  
151 ATTGTCGGCC ATTACTCCGA CCGCACTTGG AAGCCGCGTT TGGGCGGCCG  
201 CCGTCTGCCG TATCTGCTTT ATGGCACGCT GATTGCGGTT ATTGTGATGA  
251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG  
301 GCTTTGTCGT TCGGCGCGCT GATGATTGCG CTGTTAGACG TGTCGTCAAA  
351 TATGGCGATG CAGCCGTTTA AGATGATGGT CGGCGACATG GTCAACGAGG  
401 AGCAGAAAGG CTACGCCTAC GGGATTCAAA GTTCTTAGC AAATACGGGC  
451 GCGGTCGTGG CGGCGATTCT GCCGTTTGTG TTTGCGTATA TCGGTTTGGC  
501 GAACACCGCC GAGAAAGGCG TTGTGCCGCA GACCGTGGTC GTGGCGTTT  
551 ATGTGGGTGC GCGGTTGCTG GTGATTACCA GCGCGTTCAC GATTTTCAAA  
601 GTGAAGGAAT ACGATCCGGA AACCTACGCC CGTTACCACG GCATCGATGT  
651 CGCCGCGAAT CAGGAAAAAG CCAACTGGAT CGAACTCTTG AAAACCGCGC  
701 CTAAGGCGTT TTGGACGGTT ACTTTGGTGC AATTCTCTG CTGGTTCGCC  
751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGCTG  
801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACGGT  
851 ACGGCGTTTT GCGGCGCGTG CAGTCGGTTG CGGCGGTGAT TTGTTCTGTT  
901 GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTCCGGCTG  
951 TTTGGCTTTG GCGGCGCTCG GCTTTTCTC CGTTTCTTC ATCGGCAACC  
1001 AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC  
1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA  
1101 TATGGGCACT TACTTGGGCT TGTTTAAACG CTCTATCTGT ATGCCTCAAA  
1151 TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG  
1201 CAGGCCACTA TGTTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT  
1251 TTCCGTGTTC CTGATTAAAG AAACACACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 568; ORF 143>:

m143.pep  
1 MLSFGFLGVQ TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP  
51 IVGHYSDRTW KPRLGGRRLP YLLYGTIAV IVMILMPNSG SFGFGYASLA  
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKGYAY GIQSFLANTG  
151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL VITSAFTIFK  
201 VKEYDPETYA RYHGIDVAAN QEKANWIELL KTAPKAFMTV TLVQFFCWFA  
251 FQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLAHV QSVAAVICSF  
301 VLAKVPNKYH KAGYFGCLAL GALGFFSVFV IGNQYALVLS YTLIGIAWAG  
351 IITYPLTIVT NALSGKHMGT YLGLFNGSIC MPQIVASLLS FVLFPMLGGL  
401 QATMFLVGGV VLLLGAFSVF LIKEITHGGV\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m143 / g143 93.9% identity in 429 aa overlap

10 20 30 40 50 60

m143.pep	MLSFGFLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW
g143	MLSFGYLGVTQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGYSDRTW
	10 20 30 40 50 60
m143.pep	70 80 90 100 110 120
g143	KPRLGGRRLPYLLYGTLLIAVIVMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM
	70 80 90 100 110 120
m143.pep	130 140 150 160 170 180
g143	QPFKMMVGDVNEEQKYAYGIQSFLANTGAVVAAILPFVFAYIGLANTAEGVVPQTVV
	130 140 150 160 170 180
m143.pep	190 200 210 220 230 240
g143	VAFYVGAALLVITSFTIIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTA PKAFWTV
	190 200 210 220 230 240
m143.pep	250 260 270 280 290 300
g143	TLVQFFCWFAFQYMWYTYSAGAI AENVWHTTDASSVGQEAGN WYGVLA AVQSVA AVICSF
	250 260 270 280 290 300
m143.pep	310 320 330 340 350 360
g143	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT
	310 320 330 340 350 360
m143.pep	370 380 390 400 410 420
g143	NALSGKHMGTYLGLFN S I C M P Q I V A S L L S F V L F P M L G G L Q A T M F L V G G V L L L G A F S V F
	370 380 390 400 410 420
m143.pep	430
g143	LIKETHGGVX
	430

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 569>:

a143.seq	1	ATGCTCAGTT	TCGGCTTTCT	CGGCGTTCAG	ACGGCCTTTA	CCCTGCAAAG
	51	CTCGCAGATG	AGCCGCATCT	TCCAGACGCT	CGGTGCCGAT	CCGCACAGCC
	101	TCGGCTGGTT	CTTTATCCTG	CCGCCGCTGG	CGGGGATGCT	GGTGCAGCCG
	151	ATTGTGCGCC	ATTACTCCGA	CCGCACTTGG	AAGCCGCGTT	TGGGCGGCCG
	201	CCGCTGCGCG	TATCTGCTTT	ATGGCACGCT	GATTGCGGTT	ATTGTGATGA
	251	TTTGTATGCC	GAAC TCGGGC	AGCTTCGGTT	TCGGCTATGC	GTCGCTGGCG
	301	GCTTTGTGCT	TCGGCGCGCT	GATGATTGCG	CTGTTAGACG	TGTCGTCAAA
	351	TATGCGCATG	CAGCCGTTTA	AGATGATGGT	CGGCGACATG	GTCAACGAGG
	401	AGCAGAAAGG	CTACGCCTAC	GGGATTCAAA	GTTTCTTAGC	GAATACGGGC
	451	GCGGTCGTGG	CGGCGATTCT	GCCGTTTGTG	TTTGCGTATA	TCGGTTTGGC
	501	GAACACCGCC	GAGAAAGCGC	TTGTGCCGCA	GACCGTGGTC	GTGGCGTTT
	551	ATGTGGGTGC	GGCGTTGCTG	GTGATTACCA	GCGCGTTCAC	GATTTTCAAA
	601	GTGAAGGAAT	ACAATCCGGA	AACCTACGCC	CGTTACCACG	GCATCGATGT
	651	CGCCGCGAAT	CAGGAAAAAG	CCAAC TGGAT	CGAACTCTTG	AAAACCGCGC
	701	CTAAGGCGTT	TTGGACGGTT	ACTTTGGTGC	AATTCTTCTG	CTGGTTCGCG

```

751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACGTGT
851 ACGGCGTTTT GCGCGCGGTG CAGTCGGTTG CCGCGGTGAT TTGTTCTGTT
901 GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGCGGGT ATTTCGGCTG
951 TTTGGCTTTG GCGCGCTCG GCTTTTTCTC CGTTTTCTC ATCGGCAACC
1001 AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC
1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
1101 TATGGGCACT TACTTGGGCC TGTTTAACGG CTCTATCTGT ATGCCGCAA
1151 TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
1201 CAGGCCACTA TGTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT
1251 TTCCGTGTTC CTGATTAAAG AAACACACGG CGGGGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 570; ORF 143.a>:

```

a143.pep
1    MLSFGFLGVQ TAFTLQSSQM SRIFQTLGAD PHSLGWFFIL PPLAGMLVQP
51   IVGHYSDRTW KPRLGRRRLP YLLYGTIAIV IVMILMPNSG SFGFGYASLA
101  ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQGYAY GIOSFLANTG
151  AVVAAILPFV FAYIGLANTA EKVVPQTVV VAFYVGAALL VITSAFTIFK
201  VKEYNPETYA RYHGIDVAAN QEKANWIELL KTAPKAFWTV TLVQFFCWFA
251  FQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AVQSVAAVICSF
301  VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
351  IITYPLTIVT NALSGKHMGT YLGLFNGSIC MPQIVASLLS FVLFPMGLGL
401  QATMFLVGGV VLLGAFSVF LIKETHGV*

```

m143/a143 99.5% identity in 429 aa overlap

	10	20	30	40	50	60
m143.pep	MLSFGFLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW					
a143	MLSFGFLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW					
	10	20	30	40	50	60
m143.pep	70	80	90	100	110	120
a143	KPRLGRRRLPYLLYGTIAIVIMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM					
	70	80	90	100	110	120
m143.pep	130	140	150	160	170	180
a143	QPFKMMVGDMVNEEQGYAYGIGIOSFLANTGAVVAAILPFVFAYIGLANTAIEKVVPQTVV					
	130	140	150	160	170	180
m143.pep	190	200	210	220	230	240
a143	VAFYVGAALLVITSAFTIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV					
	190	200	210	220	230	240
m143.pep	250	260	270	280	290	300
a143	TLVQFFCWFAFQYMWYSAGAIENVWHTTDASSVGYQEAAGNWYGVLA AVQSVAAVICSF					
	250	260	270	280	290	300
m143.pep	310	320	330	340	350	360
a143	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT					
	310	320	330	340	350	360
m143.pep	370	380	390	400	410	420
a143	NALSGKHMGTYLGLFNGSICMPQIVASLLSFVLFPMGLGLQATMFLVGGVVLLGAFSVF					

```

a143      |||||
          NALSGKHMGTYLGLFNLSICMPQIVASLLSFVLFPMGLGLOATMFLVGGVVLLLGAFSVF
          370      380      390      400      410      420

          430
m143.pep  LIKETHGGVX
          |||||
a143      LIKETHGGVX
          430

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 571>:

```

g144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGGGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGTGC GTCTTCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTTGGC AGACGGCGTG
151 CGCGAAAACC CCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGTAGGACTT ACCGCGTGGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CGGTTTCGCAC GGGCTGGCCG TTACCCgtTT
351 CAACGCGGTG GCGGCAGACG GccgacggTt atCCCAACGA TTTGgatatT
401 TCctaccgCT TGGACGAGGA CGGCCGGCTT ACCGttaccT ATCGCGCCAC
451 CGCgctCGGC GACACGGTGT TCGACCCGAC GCTGCACATT TACTGGCGGC
501 TGGACGCGGG CCTGCACGAT GCGGTTCTGC ATATTCCGCA GGGCGGACAT
551 ATTCCGGCCG ATGCCGAAAA ACTGCCCGTC TTAACGGTTT CAGACGGCCT
601 CGAAGTATTT GA

```

This corresponds to the amino acid sequence <SEQ ID 572; ORF 144.ng>:

```

g144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTC VFVLDLGGIV QEFSVLADGV
51  RENPVVSFDD AASYADNPFO INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRRLSQR FGYFLPLGRG RPAYRYLSRH
151 RARRHGVPRD AAHLLAAGRG PARCGSAYS A GRTYSGRCRK TARLNGFRFP
201 RSI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 573>:

```

m144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGTCTGATCG ACGGGCGTGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTTGGC AGACGGCGTG
151 CGCGAAAACC TCGTGGTGTC GTTCGATGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CGGTTTCGCAC GGGCTGGCCG TTACCCGTTT
351 CAACGCGGTG GCGGCAGACG GCCGTTTCGGT GGTGCTGCGC AGCCGCTGg
401 CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTGg ATATTTCTTA
451 CCGCTTGGAC GAGGACGACC GGCTTACCGT TAaCTATCGC GCCACCGCGC
501 TCGGCGACAC GGTGTTTCGAC CCGACGCTGC ACATTTACTG GCGGCTGGAC
551 GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATGCC
601 GGCCGATGCC GAAAACTGC CCGTCTCAAC GGTTTCAGAC GACCTCGAAG
651 TATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 574; ORF 144>:

```

m144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVDLGGIV QEFSVLADGV
51  RENLVVSFDD AASYADNPFO INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLATVGRR L SQRFGFGYFL
151 PLGRGRPAYR YLSRHRARRH GVRPDAHLL AAGRGPARCG SAYSAGRTYA
201 GRCKRTARLN GFRRPRSI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m144 / g144 91.3% identity in 218 aa overlap

m144.pep	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD
g144	MSDTPATRDFGLIDGRAVTGYVLSNRRGTCVFLDLGGIVQEFSVLADGVRENVPVVSFDD
m144.pep	AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
g144	AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
m144.pep	AADGRSVVLSRLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRRHGVRPDAAHLL
g144	AAD-----GRRLSQRFG--YFLPLGRGRPAYRYLSRHRARRRHGVRPDAAHLL
m144.pep	AAGRGPARGSAYSAGRTYAGRCRKARTARLNGFRRPRISIX
g144	AAGRGPARGSAYSAGRTYSGRCRKARTARLNGFRRPRISIX

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 575>:

a144.seq

```

1   ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGTGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
101 TGGACTTGGG  CGGGATTGTG CAGGAATTTT CCGTTTGGC  AGACGGCGTG
151 CGCGAAAACC  TCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG  ATTAACAAGC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT  CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
301 GGCAGGAACG  CGCTGCACGG CGGTCGCAC  GGGCTGGCCG TTACCCGTTT
351 CAACGCGGTG  GCGGCAGACG GCCGTTCCGT GGTGCTGCGC AGCCGCTG.
401 CAACAGTCGG  CCGACGGTTA TCCCAACGAT TTGGATTGG  ATATTTCTTA
451 CCGCTTGGAC  GAGGACGACC GGCTTACCGT TACCTATCGC GCCACCGCGC
501 TCGGCGACAC  GGTGTTTCGAC CCGACGCTGC ACATTTACTG GCGGCTGGAC
551 GCGGGCCTGC  ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATTCC
601 GGCCGATGCC  GAAAAACTGC CCGTCTCAAC GGTTCAGAC  GACCTCGAAG
651 TATTGA

```

This corresponds to the amino acid sequence <SEQ ID 576; ORF 144.a>:

a144.pep

```

1   MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFSVLADGV
51  RENLVVSFDD  AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH  GLAVTRFNAV AADGRSVVLR SRLXTVGRRL SQRFGFGYFL
151 PLGRGRPAYR  YLSRHRARRH GVRPDAAHLL AAGRGPARGC SAYSAGRTYS
201 GRCRKARTARL GFRRPRSI*

```

m144/a144 99.1% identity in 218 aa overlap

m144.pep	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD
a144	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD
m144.pep	AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV

```

a144      |||||
          AASYADNPFQINKQIGRVAGRIRGAADFINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
          70      80      90      100     110     120

m144.pep  130      140      150      160      170      180
          AADGRSVVLRSLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
a144      |||||
          AADGRSVVLRSLXTVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
          130      140      150      160      170      180

m144.pep  190      200      210      219
          AAGRGPARGCSAYSAGRTYAGRCRKRTARLNGFRRPRSIX
a144      |||||
          AAGRGPARGCSAYSAGRTYSGRCRKRTARLNGFRRPRSIX
          190      200      210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 577>:

```

g146.seq
1   ATGAAGCAAA TCCCCCTCCG CCTTCTCCAG GTCGTCATTG ACCACGACAA
51  AGTCGAACAA TACGGACTGT TCGATTTTCAT GCCTTGCCCTT CGACAGCCTC
101 CTTTGGATAa ctTCCCGACT GTCCGTCCCG CGCcttTTGA GCGCGCGCGC
151 AAGCACGTCG AAAGAAGGCG GCAGGATAAA GATACCGACA GCTTCCGGCA
201 GCGCGTTGCG AACCTGCGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAGCCTG CCGCCGCCAA CGCATTACAG CCCTCCGTGC TTGTGCCGTA
301 ATAGTTGCCG AATACGTCG CGTATTCCAA AAAAGCCTCC TGCGCGATAA
351 GCGATTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGGCG CGTCGTATGC GACACGGAAG CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TTCGATAAGC GGACATATTT
551 TTTACCTGTA TATTTTCCAA CCGATTGTAT CACAACGGAC ACCCTATTTT
601 ATATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 578; ORF 146.ng>:

```

g146.pep
1   MKQIPLRLQ VVIDHDKVEQ YGLDFMPCL RQPPLDNFPT VRPAPFEARG
51  KHVERRRQDK DTDSFRQVA NLRRALNVDF QNHVIACRRQ RIHALRACAV
101 IVAEYVCVFQ KSLLRDKREF LFFGNKVIMY AVCFAFTRRA RMRHGNAQT
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQRTPYF
201 IFADAHILPL LF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 579>:

```

m146.seq
1   ATGGCGCAAA TCCTCCTCCG CTCGCGCCAA GTCGTCATTG ACCACGACAA
51  AGTCAAACAA TACGGACTGC TCGATTTTCAT GCCTTGCCCTT CGACAGCCTC
101 CTTTGGATAA CTTCCCGACT GTCCGTCCCG CGTCCGTTGA GCGCGCGCGC
151 AAGTACGTCG AAAGAAGGCG GCAGGATAAA GATGCCGACG GCTTCCGGCA
201 GCGCGTCGCG AACCTGCGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAGCCTG CCGCCGCCAA CGCATTACAG CCCTCCGTGC TTGTGCCGTA
301 ATAGTTGCCA AATACGTCG CGTATTCCAA AAAAGCCTCC TGCGCGATAA
351 GCGACTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGGCG CGTCGTGTGC GACACGGAAG CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TTCGATAAGC GGACATATTT
551 TTTACCTGTA TATTTTCCAG CCGATTGTAT CACAATGGAC ACCCAGTTTC
601 CTATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 580; ORF 146>:

```

m146.pep
1   MAQILLRSRQ VVIDHDKVKQ YGLDFMPCL RQPPLDNFPT VRPASVEARG
51  KYVERRRQDK DADGFGQVA NLRRALNVDF QNHVIACRRQ RIHTLRACAV
101 IVAKYVGVFQ KSFLRDKRLK LFFGNKVIMY AVCFAFTRRA RVRHGNAQT
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQWTPSF
201 LFADAHILPL LF*

```



		10	20	30	40	50	60
m146.pep		MAQILLRSRQVV	IDHDKVKQYGL	LDLDFMPCLRQ	PPLDNFPTVR	PASVEARGKYV	ERRRQDK
g146		MKQIPLRLLQV	IDHDKVEQYGL	FDLDFMPCLRQ	PPLDNFPTVR	PAPFEARGKHV	ERRRQDK
		10	20	30	40	50	60
		70	80	90	100	110	120
m146.pep		DADGFGQRVAN	LRRLNVDFQNH	VIACRRQRIHT	LRACAVIVAKY	VGVFQKSFLR	DKRRLK
		:					
g146		DTDSFRQRVAN	LRRLNVDFQNH	VIACRRQRIHA	LRACAVIVAEY	VCVFQKSLLR	DKRFRK
		70	80	90	100	110	120
		130	140	150	160	170	180
m146.pep		LFFGNKVMIMY	AVCFAFTRRARR	VRHGNATVMVC	QQRHQRGFAR	AGSGRNDKDV	AFSIS
g146		LFFGNKVMIMY	AVCFAFTRRARR	MRHGNATVMVC	QQRHQRGFAR	AGSGRNDKDV	AFSIS
		130	140	150	160	170	180
		190	200	210			
m146.pep		GHIFYLYIFQPI	VSQWTPSFLFAD	AHILPLLFX			
g146		GHIFYLYIFQPI	VSQRTPYFIFAD	AHILPLLFX			
		190	200	210			

```

al46.seq
1  ATGGCGCAAA TCCTCCTCCG CCCGCGCCAA GTCATCATTG ACCACGACAA
51  AATCGAACAA TACGGACTGT TCGATTTTCAT GCCTTGCCCTT CGACGACCTC
101 CTTTGGAATA CTTCGCCGACT GTCCGTCGCCG CGTCCGTGGA GACGCGCAGC
151 AAGCACATCG AAAGACGCGCG GCAGGATAAAA GATGCCGAGG GCTTCGGGGA
201 GCGCATCTCG AACCTGAGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAACCTG CCGCCGCGCAA CGCATTCACA CCCTCCGCGC TTGTGCCGTA
301 ATAGTTGCCG AACACGTCCG CGTATTCCAA AAAAGCCTCC TCGCGGATAA
351 GCGACTCAAA CTCTTCTTTG GAAACAAAGT GACAAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGACG CGTCGTGTGC GACACGGAAA CCGCGAAACC
451 GTTATGTTTT GCCAACAGCC GCGACACCAG CGTGGTTTTG CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCCTT TTCGATAAGC GGACATATTT
551 TTTACCTGTA TATTTTCCAG CCGATTGTAT CACAACGGAC ACCCGGTTTC
601 CTATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTTTGA

```

a146.pep

1	MAQILLRPRQ	VIIDHDKIEQ	YGLDFEMPCL	RQPPLDNFPT	VRPASVETRS
51	KHIERRRQDK	DADGFGQRIS	NLSRALNVDF	QNHVITCRRO	RIHTLRACAV
101	IVAEHRVRFQ	KSLLRDKRLK	LFFGNKVMIM	AVCFATRRT	RRVRHGNAQT
151	VMVCQQPRHQ	RGFARAGSGR	NDKDVAFSIS	GHIFYLYIFQ	PIVSQRTPGF
201	LEADAHILPL	LF*			

	10	20	30	40	50	60
m146.pep	MAQILLRSRQVVIDHDKVKQYGLLDFMPCLRPPLDNFPTVRPASVEARGKYVERRRQDK					
		:       :     :			:     :     :	
a146	MAQILLRPRQVIIDHDKIEQYGLFDMPCLRPPLDNFPTVRPASVETRSKHIERRRQDK					
	10	20	30	40	50	60
	70	80	90	100	110	120

```

m146.pep      DADGFGQRVANLRRALNVDFQNHVIACRRQRIHTLRACAVIVAKYVGVFQKSFLRDKRLK
                |||::| |||||::| |||||::| |||||::| |||||::| |||||::|
a146          DADGFGQRI SNLSRALNVDFQNHVITCRQRRIHTLRACAVIVAEHVRVFQKSLLRDKRLK
                70          80          90          100         110         120

                130          140          150          160          170          180
m146.pep      LFFGNKVIMYAVCFATRRRARRVRHGNQATVMVCQQPRHQRGFARAGSGRNDKDVAFSIS
                |||::| |||||::| |||||::| |||||::| |||||::| |||||::| |||||::|
a146          LFFGNKVIMYAVCFATRRRTRVRHGNQATVMVCQQPRHQRGFARAGSGRNDKDVAFSIS
                130          140          150          160          170          180

                190          200          210
m146.pep      GHIFYLYIFQPIVSQWTPSFLFADAHILPLFLX
                |||::| |||||::| |||||::| |||||::| |||||::| |||||::|
a146          GHIFYLYIFQPIVSQRTPGFLFADAHILPLFLX
                190          200          210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 583>:

g147.seq (partial)

```
1 ..ATGCGACGAG AAGCCAAAAT GGCACAAATC AACTCAAAC CCATTGTTTT
51 ATCAATTCTT TTAATCAACA CACCCCTCCT CGCCCAAGCG CATGAAACTG
101 AGCAATCGGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCCG
151 CGCGCGACTT CGGGCTGCT GCACACTTCG ACCGCCTCCG ACAAAATCAT
201 CTCCGGCGAT ACTTTGCGCC AAAAAGCCCT CAACTTGGCG GACGCTTTGG
251 ACGGCGTACC GGGCATCCAC GCTTCGCAAT ACGCGGCGG CGCATCOGT
301 CCCGTTATTC GCGGTCAAAC GGGCAGACGG ATTAAGTAT TGAACCATCA
351 CGGCGAAACG GGCGATATGG CGGACTTTTC TCCCGATCAC GCCATTATGG
401 TAGATACCGC CTTGTGCGAA caggttGAAA TCCTGCGCGG GCCGGTTACG
451 CTCTTGATACA GCTCGGgcaa tgtggccgGG GCTGGtcaat gttgccgatg
501 gAAAAAtccc ccaaaaaAAtg cc..
```

This corresponds to the amino acid sequence <SEQ ID 584; ORF 147.ng>:

g147.pep (partial)

1	..MRREAKMAQI	TLKPIVLSIL	LINTPLLAQA	HETEQSVGLE	TVSVVGKSRP
51	RATSGLLHTS	TASDKIISGD	TLRQKAVNLG	DALDGVPGIH	ASQYGGGASA
101	PVIRGGQTGR	IKVLNHHGET	GDMADFSPPH	AIMVDTALSQ	QVEILRPVPT
151	LLYSGTNVAG	AGQCCRWKNP	PKNA..		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 585>:

```

m147.seq (partial)
1 ..CCGCATAAAA CTGAGCAATC GGTGGATTTC GAAACGGTCA GCGTCGTCGG
51 CAAAAGCCGT CCGCGCGCCA CGTCGGGGCT GTTGACACT TCGACCGCCT
101 CCGACAAAT CATCTCCGG GATACCTTGC GCCAAAAGC CGTCAACTGG
151 GGCAGCGCTT TAGACGGCCT ACCGGGCAT CACGCTTCG ATACACGGCG
201 CGGCGCGTCT GCTCCCGTCA TTCGCGGTCA AACAGGCAGG CGGATTAAAG
251 TGTTGAACCA TCACGGCGAA ACAGGCGATA TGGCGGATTT TTCGCCGAT
301 CACGCCATTA TGGTAGATAC CGCCTTGTC CAACAGGTCG AAATCCTGCG
351 CGGGCCGGTT ACGCTCTGT ACAGCTCGGG CAATGTGGC GGGCTGGTCG
401 ATGTTGCCGA TGGCAAAATC CCCGAAAAAA TGCCTGAAAA CGCGGTATCG
451 GGCGAACTCG GATTGCGTTT GAGCAGCGGC AATCTGAAAA AACTCACGTC
501 CGGCGGCATC AATATCGGTT TGGGCAAAAA CTTTGTATTG CACACGGAAG
551 GGCTGTACCG CAAATCGGGG GATTACGCC TACCCGCTTA CGCAATCTG
601 AAACGCCTGC CCGACAGCCA GCCCGATTG CAAACGGGCA CGCTCGGGCT
651 GTCTTGGGTT GGCGAAAAAG GTTTTATCGG CGTAGCGTAC AGCGACCGTC
701 GCGACCAATA TGGTCTGCCT GCCCAGACC ACGAATACGA TGATTGCCAC
751 GCGGACATCA TCTGGCAAAA GAGCTTGATT ACAAACCGCT ATTTACAGT
801 TTATCCGCAC CTGTTGACCG AAGAAGACAT GATTACGAC AATCCGGGCT
851 TGAGCTGCGG CTTCACGAC GACGATAATG CACACGCACA CACCCACAGC
901 GGCAGACCGT GGATAGACCT GCGCAACAAA CGCTACGAAC TCCGTGCCGA
951 ATGGAAGCAA CCGTTCCCCG GTTTTGAAGC CCTGCGCGTA CGACTGAACC
1001 GCAACGACTA CCGCCACGAC GAAAAAGCAG CGATGCGAGT CACAAAATTT
1051 TTAAACAACC AAACGCAAAA CGCCCGCATC GAGTTGCGCC ACCAACCCTT

```

```

1101 AGGTCGTCTG AAAGGCAGCT GGGGCGTGCA ATATTACAA CAAAAATCCA
1151 GTGCTTTATC TGCCATATCC GAAGCGGTTA AACAAACGAT GCTGCTTGAC
1201 AACAAAGTGC AACATTACAG CTTTTCGGT GTAGAACAGG CAAACTGGGA
1251 CAACTTCACG CTTGAAGGAG GCGTACGCGT GGAAAAACAA AAAGCCTCCA
1301 TTCAGTACGA CAAAGCATTG ATTGATCGGG AAAACTACTA CAACCACCCC
1351 CTGCCCCGACC TCGGCGCGCA CCGCCAAACC GCCCGTCAT TCGCACTTTC
1401 GGGCAACTGG TATTTACGCG CACAACACAA ACTCAGCCTG ACCGCCTCCC
1451 ATCAGGAACG CCTGCCGTCA ACGCAAGAGC TGTACGCACA CGGCAAACAC
1501 GTCGCCACCA ACACCTTTGA AGTCGGCAAC AAACACCTCA ACAAAGAGCG
1551 TTCCAACAAT ATCGAACTCG CGCTGGGCTA CGAAGGCGAC CGCTGGCAAT
1601 ACAATCTGGC ACTCTACCGC AACCGCTTCG GTAACACAT TTACGCCCAA
1651 ACCTTAAACG ACGGACGCGG CCCCAAATCC ATCGAAGACG ACAGCGAAAT
1701 GAAGCTCGTG CGCTACAACC AATCCGGCGC CGACTTCTAC GGCGCGGAAG
1751 GCGAAATCTA CTTCAAACCG ACACCGCGCT ACCGCATCGG CGTTTCCGGC
1801 GACTATGTAC GAGGCGGTCT GAAAAACCTG CCTTCCCTAC CCGGCAGAGA
1851 AGATGCCTAC GGCAACCGTC CTTTCATCGC ACAGGACGAC CAAAATGCCC
1901 CCCGTGTTCC GGCTGCGCGC CTCGGCTTCC ACCTGAAAGC CTCGTGACC
1951 GACCGTATCG ATGCCAATT GGACTACTAC CGCGTGTTCC CCCAAAAACA
2001 ACTCGCCCGC TACGAAACGC GCACGCGCGG ACACCATATG CTCAACCTCG
2051 GCGCAACTA CCGCCGCAAT ACGCGCTATG GCGAGTGGAA TTGGTACGTC
2101 AAAGCCGACA ACCTGCTCAA CCAATCCGTT TACGCCACA GCAGCTTCT
2151 CTCTGATACG CCGCAAATGG GCCGCAGCTT TACCGGCGGC GTGAACGTGA
2201 AGTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 586; ORF 147>:

```

m147.pep (partial)
1 ..PHKTEQSVDL ETVSVVGKSR PRATSGLLHT STASDKIISG DTLRQKAVNL
51 GDALDGVPGI HASQYGGGAS APVIRGQTGR RIKVLNHHGE TGMADFSPPD
101 HAIMVDTALS QQVEILRGPV TLLYSSGNVA GLVDVADGKI PEKMPENGVS
151 GELGLRLSSG NLEKLTSGGI NIGLGKNFVL HTEGLYRKSG DYAVPRYRNL
201 KRLPDSHADS QTGSIGLSWV GEKGFIVGAY SDRRDQYGLP AHSHEYDDCH
251 ADIIWQKSLI NKRYLQLYPH LLTEEDIDYD NPGLSCGFHD DDNAHAHTHS
301 GRPWIDLRNK RYELRAEWKQ PFPGFELRV HLNRRNDYRD EKAGDAVENF
351 FNNQTONARI ELRHQPIGRL KGSWGVQYLQ QKSSALS AIS EAVKQPMLLD
401 NKVQHYSFFG VEQANWDNFT LEGGVRVEKQ KASIYQDKAL IDRENYNHP
451 LPDLGAHRQT ARSFALSGNW YFTPOHKL SL TASHQERLPS TQELYAHGKH
501 VATNTFEVGN KHLNKERSNN IELALGYEGD RWQYNLALYR NRGNYIY AQ
551 TLNDGRGPKS IEDDSEM KLV RYNQSGADFY GAEGEYIFKP TPRYRIGVSG
601 DYVRGRLKNL PSLPGREDAY GNRPFIAQDD QNAPRVPAAR LGFHLKASLT
651 DRIDANLDYY RVFAQNKLAR YETRTPGHHM LNLGANYRRN TRYGEWNWYV
701 KADNLLNQSV YAHSSFLSDT PQMGRSFTGG VNVKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m147 / g147 92.3% identity in 142 aa overlap

```

m147.pep                                10      20      30
                                         PHKTEQSVDL ETVSVVGKSR PRATSGLLHTS
g147                                10      20      30      40      50      60
MRREAKMAQITLKP IVLSILLINTPLLAQA HETE QSVGLETVSVVGKSR PRATSGLLHTS

m147.pep                                40      50      60      70      80      90
TASDKIISGDTLRQKAVNLGDALDGVPGI HASQYGGGASAPVIRGQTGRRIKVLNHHGET
g147                                40      50      60      70      80      90      100      110      120
TASDKIISGDTLRQKAVNLGDALDGVPGI HASQYGGGASAPVIRGQTGRRIKVLNHHGET

m147.pep                                100      110      120      130      140      150
GDMADFSPPDHAIMVDTALSQQVEILRGPV TLLYSSGNVAGLVDVADGKIPEKMPENGVS
g147                                100      110      120      130      140      150
GDMADFSPPDHAIMVDTALSQQVEILRGPV TLLYSSGNVAGAGQCCRWKNPKNA

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 587>:

a147.seq

```

1  ATGCGACGAG AAGCCAAAAT GGCACAAACT ACACTCAAAC CCATTGTTTT
51  ATCAATTCTT TTAATCAACA CACCCCTCCT CTCCCAAGCG CATGGAAGCTG
101 AGCAATCAGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCCG
151 CGCGCCACTT CGGGGCTGCT GCACACTTCT ACCGCCTCCG ACAAATCAT
201 CAGCGGCGAC ACCTTGCGAC AAAAAGCCGT CAACTTGGGT GATGCTTTAG
251 ACGGCGTACC GGGCATTTCAT GCCTCGCAAT ACGGCGGCGG CGCATCCGCT
301 CCCGTTATTC GCGGTCAAAC AGGCAGACGG ATTAAAGTGT TGAACCATCA
351 CGGCGAAACG GCGGACATGG CGGACTTCTC TCCAGACCAT GCAATCATGG
401 TGGACAGCGC CTGTGTCGAA CAGGTCGAAA TCCTGCGCGG TCCGGTTACG
451 CTCTTGTAACA GCTCGGGCAA TGTGGCGGGG CTGGTCGATG TTGCCGATGG
501 CAAAATCCCC GAAAAAATGC CTGAAAACGG CGTATCGGGC GAACTCGGAT
551 TCGGTTTGAG CAGCGGCAAT CTGGAAAAAC TCACGTCCGG CGGCATCAAT
601 ATCGGTTTGG GCAAAAACCT TGTATTGCAC ACGGAAGGGC TGTACCGCAA
651 ATCGGGGGAT TACGCCGTAC CGCGTTACCG CAATCTGAAA CGCCTGCCCG
701 ACAGCCACGC CGATTTCGAA ACGGGCAGCA TCGGGCTGTC TTGGGTTGGC
751 GAAAAAGGCT TTATCGGCGC AGCATAACAG GACCGTCGCG ACCAATATGG
801 TCTGCCTGCC CACAGCCACG AATACGATGA TTGCCACGCC GACATCATCT
851 GGCAAAAGAG TTTGATTAAC AAACGCTATT TGCAGCTTTA TCCGCACCTG
901 TTGACCGAAG AAGACATCGA TTACGACAAT CCGGGCTTGA GCTGCGGCTT
951 TCACGACGAC GATGATGCAC ACGCCCATGC CCACAACGGC AAACCTTGGA
1001 TAGACCTGCG CAACAAACGC TACGAACTCC GCGCCGAATG GAAGCAACCG
1051 TTCCCGGTTT TTGAAGCCCT GCGCGTACAC CTGAACCGCA ACGACTACCG
1101 CCACGACGAA AAAGCAGGCG ATGCAGTAGA AAACCTTTTT AACAACCAAA
1151 CGCAAAACGC CCGTATCGAG TTGCGCCACC AACCCATAGG CCGTCTGAAA
1201 GGCAGCTGGG GCGTGCAATA TTTGGGACAA AAATCCAGTG CTTTATCTGC
1251 CACATCCGAA GCGGTCAAAC AACCAGTCTG GCTTGACAAT AAAGTGCAAC
1301 ATTACAGCTT TTTCCGGTGT GAACAGGCAA ACTGGGACAA CTTACGCTT
1351 GAAGGCGGCG TACGCGTGGA AAAACAAAAA GCCTCCATCC GCTACGACAA
1401 AGCATTGATT GATCGGGAAA ACTACTACAA CCATCCCCTG CCCGACCTCG
1451 GCGCGCACCG CCAAACCGCC CGCTCATTTC CACTTTCGGG CAACTGGTAT
1501 TTCACGCCAC AACACAAACT CAGCCTGACC GCCTCCATC AGGAACGCCT
1551 GCGGTCAACG CAAGAGCTGT ACGCACACGG CAAACACGTC GCCACCAACA
1601 CCTTTGAAGT CGGCAACAAA CACCTCAACA AAGAGCGTTC CAACAATATC
1651 GAACTCGCGC TGGGCTACGA AGGCGACCGC TGGCAATACA ATCTGGCACT
1701 CTACCGCAAC CGCTTCGGCA ACTACATTTA CGCCCAAACC TTAACGACG
1751 GACGCGGCCC CAAATCCATC GAAGACGACA GCGAAATGAA GCTCGTGCGC
1801 TACAACCAAT CCGGTGCGGA CTTCTACGGC GCGGAAGGCG AAATCTACTT
1851 CAAACCGACA CCGCGCTACC GCATCGGCGT TTCCGGCGAC TATGTACGAG
1901 CCGGTCTGAA AAACCTGCCT TCCCTACCCG GCAGGGAAGA CGCCTACGGC
1951 AACCGCCCAC TCATTGCCCA AGCCGACCAA AACGCCCTC GCGTTCGGCG
2001 TGCGCGCCTC GGCGTCCACC TGAAAGCCTC GCTGACCGAC CGCATCGATG
2051 CCAATTTGGA CTACTACCGC GTGTTCGCCC AAAACAAACT CGCCCGCTAC
2101 GAAACGCGCA CGCCCGGACA CCATATGCTC AACCTCGGCG CAAACTACCG
2151 CCGCAATACG CGCTATGGCG AGTGGAATTG GTACGTCAA GCGGACAACC
2201 TGCTCAACCA ATCCGTTTAC GCCACAGCA GCTTCCTCTC TGATACGCCG
2251 CAAATGGGCC GCAGCTTTAC CGGCGGCGTG AACGTGAAGT TTTAA

```

This corresponds to the amino acid sequence <SEQ ID 588; ORF 147.a>:

a147.pep

```

1  MRREAKMAQT TLKPIVLSIL LINTPLLSQA HGTEQSVGLE TVSVVGKSRP
51  RATSGLLHTS TASDKIISGD TLRQKAVNLG DALDGVPGIH ASQYGGGASA
101 PVIRGQTGRR IKVLNHHGET GDMADFSPDH AIMVDSALSQ QVEILRGVPT
151 LLYSSGNVAG LVDVADGKIP EKMPENGVSQ ELGLRLSSGN LEKLTSGGIN
201 IGLGKNFVLH TEGLYRKSGD YAVPRYRNK RLPDSHADSQ TGSIGLSWVG
251 EKGFIGAAYS DRRDQYGLPA HSHEYDDCHA DIIWQKSLIN KRYLQLYPHL
301 LTEEDIDYDN PGLSCGFHDD DDAHAHAHNG KPWIDLNRKR YELRAEWKQP
351 FPGFEALRVH LNRNDYRHDE KAGDAVENFF NNQTONARIE LRHQPIGRK
401 GSWGVOYLQ KSSALSATSE AVKQPMLLDN KVQHSFFGV EQANWDNFTL
451 EGGVRVEKQK ASIRYDKALI DRENYNHNPL PDLGAHRQTA RSFALSGNWX
501 FTPQHKLST ASHQERLPST QELYAHGKHV ATNTFEVGNK HLNKERSNNI
551 ELALGYEGDR WQYNLALYRN RFGNYIYAQT LNDGRGPKSI EDDSEMKLVR

```

601 YNQSGADFYG AEGEIYFKPT PRYRIGVSGD YVRGRLKNLP SLPGREDAYG  
651 NRPLIAQADQ NAPRVPAARL GVHLKASLTD RIDANLDYYR VFAQNKLARY  
701 ETRTPGHHML NLGANYRRNT RYGEWNWYVK ADNLLNQSVY AHSSFLSDTP  
751 QMGRSFTGGV NVKF\*

m147/a147 98.1% identity in 734 aa overlap

```

                                10      20      30
m147.pep                      PHKTEQSV DLETVSVVGKSRPRATSGLLHTS
                                | | | | | | | | | | | | | | | | | |
a147                          MRREAKMAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLETVSVVGKSRPRATSGLLHTS
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m147.pep                      TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                                | | | | | | | | | | | | | | | | | |
a147                          TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                                70      80      90      100     110     120

                                100     110     120     130     140     150
m147.pep                      GDMADFSPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVDVADGKIPEKMPENGVS
                                | | | | | | | | | | | | | | | | | |
a147                          GDMADFSPDHAIMVDSALSQQVEILRGPVTLTYSSGNVAGLVDVADGKIPEKMPENGVS
                                130     140     150     160     170     180

                                160     170     180     190     200     210
m147.pep                      ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQ
                                | | | | | | | | | | | | | | | | | |
a147                          ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQ
                                190     200     210     220     230     240

                                220     230     240     250     260     270
m147.pep                      TGSIGLSWVGEGKGFIVAYSDDRDQYGLPAHSHEYDDCHADI IWQKSLINKRYLQLYPHL
                                | | | | | | | | | | | | | | | | | |
a147                          TGSIGLSWVGEGKGFIAAYSDDRDQYGLPAHSHEYDDCHADI IWQKSLINKRYLQLYPHL
                                250     260     270     280     290     300

                                280     290     300     310     320     330
m147.pep                      LTEEDIDYDNPGLSGCFHDDDDNAHAHATHSGRPWIDLNRKRYELRAEWKQFPFGFEALRVH
                                | | | | | | | | | | | | | | | | | |
a147                          LTEEDIDYDNPGLSGCFHDDDDAHAHANGKRPWIDLNRKRYELRAEWKQFPFGFEALRVH
                                310     320     330     340     350     360

                                340     350     360     370     380     390
m147.pep                      LNRNDYRHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLGQKSSALSATSE
                                | | | | | | | | | | | | | | | | | |
a147                          LNRNDYRHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLGQKSSALSATSE
                                370     380     390     400     410     420

                                400     410     420     430     440     450
m147.pep                      AVKQPMLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASI QYDKALIDRENYNHPL
                                | | | | | | | | | | | | | | | | | |
a147                          AVKQPMLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPL
                                430     440     450     460     470     480

                                460     470     480     490     500     510
m147.pep                      PDLGAHRQTARSAFALSGNWFYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK
                                | | | | | | | | | | | | | | | | | |
a147                          PDLGAHRQTARSAFALSGNWFYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK
                                490     500     510     520     530     540

                                520     530     540     550     560     570
m147.pep                      HLNKERSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR
                                | | | | | | | | | | | | | | | | | |
a147                          HLNKERSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR
```

	550	560	570	580	590	600
m147.pep	580	590	600	610	620	630
	YNQSGADFYGAEGEIFYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFFIAQDDQ					
a147	YNQSGADFYGAEGEIFYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPPLIAQADQ					
	610	620	630	640	650	660
m147.pep	640	650	660	670	680	690
	NAPRVPAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNT					
a147	NAPRVPAARLGVHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNT					
	670	680	690	700	710	720
m147.pep	700	710	720	730		
	RYGEWNWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
a147	RYGEWNWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
	730	740	750	760		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 589>:

g148.seq

1	ATGGCGTTAA	AAACATCAAA	CTTGGAACAC	GCAATGctgg	ttcaTCCCGA
51	AgctATgagt	gtcggcgCGC	TTGccgAcaa	AATCCGCAAA	AtcgaAAact
101	gGCCGCAAAA	AGgcaTCTTA	TTCCACGACA	TCACGCCCGT	CCTGCAAAGT
151	GCGGAATACT	TCCGCCTTTT	GGTCGATTTG	CTGGTTTACC	GCTATATGGA
201	TCAGAAAATC	GACATCGTTG	CCGGCTTGGA	CGCGCGCGGC	TTCATTATCG
251	GCGCGGCACT	CGCCTACCAG	CTCAaCGtcg	gctTCGTCCC	CATCCGCAAA
301	AAAGGCAAGC	TGCCTTTTGA	AACCGTATCG	CAAAGCTAcg	cgctCGAATA
351	CGGGGAAGCT	GCGGTGGAAA	TCCACACCGa	tgccgTCAAA	CCCGGTTTCGC
401	GCGTCTGCT	GGTCGATGAT	TTGGTTGCCA	CGGGCGGCAC	AATGCTTGCC
451	GGGCTGGAAC	TGATCCGCAA	ACTCGGCGGG	GAAATTGTCT	AAgcccgcgC
501	CATTTTGGAA	TTTACCGACC	TTCAAGGCGG	CAAGAATATC	CGCGCAAGTG
551	GCGCGCCCTT	ATTTACCCTG	CTTCAAACG	AAGGCTGCAT	GAAAGGCTGA

This corresponds to the amino acid sequence <SEQ ID 590; ORF 148.ng>:

g148.pep

1	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQS
51	AEYFRLLVDL	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVGFPPIRK
101	KGKLPFETVS	QSYALEYGEA	AVEIHTDAVK	PGSRVLLVDD	LVATGTMILA
151	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI	RASGAPLFTL	LQNEGCMKG*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 591>:

m148.seq

1	ATGGCGTTAA	AAACATCAAA	CTTGGAACAC	GCAATGCTGG	TTCATCCCGA
51	AGCTATGAGT	GTCGGCGCGC	TTGCCGACAA	AATCCGCAAA	ATCGAAAAC
101	GGCCGCAAAA	AGGCATCTTA	TTCCACGACA	TCACGCCCGT	CCTTCAAAGC
151	GCGGAATACT	TCCGCCTTTT	GGTTGATTTA	TTGGTTTACC	GCTATATGGA
201	TCAGAAAATC	GACATCGTTG	CCGGTTTGGA	CGCGCGCGGC	TTCATTATCG
251	GCGCGGCACT	CGCCTACCAG	CTCAACGTCG	GTTCGTCCC	CATCCGCAAA
301	AAAGGCAAGC	TGCCTTTTGA	AACCGTATCG	CAAAGCTACG	CGCTCGAATA
351	CGGGGAAGCT	GCGGTGGAAA	TCCACACCGA	TGCCGTCAAA	CTCGGTTTCGC
401	GCGTCTGCT	GGTCGATGAT	TTGATTGCCA	CGGGCGGCAC	GATGCTTGCC
451	GGACTGGAAC	TGATCCGCAA	ACTCGGCGGA	GAAATTGTCT	AAGCCGCCGC
501	CATTTTGGAA	TTTACCGACC	TTCAAGGCGG	CAAGAATATC	CGTGCAAGCG
551	GCGCGCCCTT	ATTTACCCTG	CTTCAAACG	AAGGCTGTAT	GAAGGCGTGA

This corresponds to the amino acid sequence <SEQ ID 592; ORF 148>:

m148.pep

1	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQS
51	AEYFRLLVDL	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVGFPPIRK

101 KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LIATGGTMLA  
 151 GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m148 / g148 99.0% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
g148	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m148.pep	LVYRYMDQKIDIVAGLDARGFIIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA					
g148	LVYRYMDQKIDIVAGLDARGFIIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m148.pep	AVEIHTDAVKLGSRVLLVDDLIATGGTMLAGLELIRKLGG EIVEAAAILEFTDLQGGKNI					
g148	AVEIHTDAVKPGSRVLLVDDLVATGGTMLAGLELIRKLGG EIVEAAAILEFTDLQGGKNI					
	130	140	150	160	170	180
	190	200				
m148.pep	RASGAPLFTLLQNEGCMKGX					
g148	RASGAPLFTLLQNEGCMKGX					
	190	200				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 593>:

a148.seq

1	ATGGCGTTAA	AAACATCAAA	CTTGGAACAC	GCAATGCTGG	TTCATCCCGA
51	AGCTATGAGT	GTCGGTGCGC	TTGCCGACAA	AATCCGCAAA	ATCGAAAAC
101	GGCCGCAAAA	AGGCATCTTA	TTCCACGACA	TCACGCCCGT	CCTGCAAAGC
151	GCGGAATACT	TCCGACTTTT	GGTTGATTTA	TTGGTTTACC	GCTATATGGA
201	TCAGAAAATC	GACATCGTTG	CCGGTTTGGA	CGCGCGCGGC	TTCATTATCG
251	GCGCGGCACT	CGCCTACCAG	CTCAACGTCG	GTTTCGTCCC	CATCCGCAAA
301	AAAGGCAAGC	TGCCTTTTGA	AACCGTATCG	CAAAGCTACG	CGCTCGAATA
351	CGGGGAAGCT	GCGGTGGAAA	TCCACACCGA	TGCCGTCAAA	CTCGGTTTCG
401	GCGTGCTGCT	GGTCGATGAT	TTGGTTGCCA	CGGGCGGCAC	GATGCTTGCC
451	GGACTGGAGC	TGATCCGCAA	ACTCGGCGGG	GAAATTGTCT	AAGCCGCGCG
501	CATTTTGGAA	TTTACCAGAC	TTCAAGGCGG	CAAGAATATC	CGTGCAAGCG
551	GCGCGCCCTT	ATTTACCCTG	CTTCAAAACG	AAGGCTGTAT	GAAGGGCTGA

This corresponds to the amino acid sequence <SEQ ID 594; ORF 148.a>:

a148.pep

1	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQS
51	AEYFRLLVDL	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVG FVPIRK
101	KGKLPFETVS	QSYALEYGEA	AVEIHTDAVK	LGSRVLLVDD	LVATGGTMLA
151	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI	RASGAPLFTL	LQNEGCMKG*

m148/a148 99.5% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
a148	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
	10	20	30	40	50	60
	70	80	90	100	110	120

```

m148.pep  LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
a148       LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
           70      80      90      100     110     120

           130     140     150     160     170     180
m148.pep  AVEIHTDAVKLGSRVLLVDDLIATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
a148       AVEIHTDAVKLGSRVLLVDDLVATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
           130     140     150     160     170     180

           190     200
m148.pep  RASGAPLFTLLQNEGCMKGX
a148       RASGAPLFTLLQNEGCMKGX
           190     200

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 595>:

```

g149.seq
1  ATGTTGATTG ACAACAATGT CCGCCATTAC AGCTTTTTCG GTGTAGAACA
51  GGCAAATTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATCCGGTAC GACAAAGCAT TGATTGATCG AGAAAACTAC
151 TACAACCAGC CCCTGCCCCG CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 GTTCGCACTT TCGGGCAACT GGTATTTCAC GCCACACCAC AAACCTCAGCC
251 TGACCGCCTC CCATCAGGAA cgCCTGCCGT CAACGCaaga actGtACgca
301 cacggcAAGC ACgtcgccac CAACACCTTT GAagtggca acaaACACCT
351 CAACAAAGaG Cgttccaaca atatcgaACT CGCGCTGGGc tAcaaaggcg
401 accGCTGGCA ATACAATCTG GCAGCCTACC GCAACCGAtT CGGCAACTAC
451 ATTTACGCCC AAACCTTAaa cgacggacgC GGCCCCAAAT CCATCgaaga
501 cgacagcgaa ATGaagctCG TCGCTACAA CCAATCCGGT GCCGACTTCT
551 ACGgcgcgga aggcgaaatc tACTTcaaaC CGAcACCGCG CTACCGCATC
601 GGTGTTTCCG GCGACTatgt acgaggccgT CTGAAAAACC TGCCGTCCCT
651 ACCCGGCAGG gaagatccCT AcggcAAACG TCccttcaTC GCACAAGCCG
701 ACCAAAACGC CCCCAGCATT cgggtGCGC GCCTCGGCTT CCACCTGAAA
751 ACCTCGCTAA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCCGGTGTT
801 CGCCCCAAAC AAACCTGCCC GCTACGAAAC GCGTACGCCC GGACACCATA
851 TGCTCAACCT CCGTGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGctc aACcaatCcg tTTACGCCCa
951 CAGCAGCTC CTCTCTGATA CGCCGCAAAt gGGCCGACG TTTgccgGCG
1001 gcgtaAACGT GaAGTTtaa

```

This corresponds to the amino acid sequence <SEQ ID 596; ORF 149.ng>:

```

g149.pep
1  MLIDNNVRHY SFFGVEQANW DNFTLEGGVR VEKOKASIRY DKALIDRENY
51  YNQPLPDLGA HRQTARSFAL SGNWYFTPHH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YKGDWRQYNL AAYRNRFGNY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVS GDYVRGR LKNLPSLPGR EDPYGKRPF I AQADQNAPRI PAARLGFHLK
251 TSLTDRIDAN LDYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NOSVYAHSSF LSDTPQMGRS FAGGVNVKF*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 597>:

```

m149.seq
1  ATGCTGCTTG ACAACAAAGT GCAACATTAC AGCTTTTTTCG GTGTAGAACA
51  GGCAAACCTG GACAACTTCA CGCTTGAAGG AGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATTAGTAC GACAAAGCAT TGATTGATCG GGAAAACTAC
151 TACAACCACC CCCTGCCCCG CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 ATTCGCACTT TCGGGCAACT GGTATTTCAC GCCACAACAC AAACCTCAGCC
251 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCaAGA GCTGTACGCA
301 CACGGCAAAC ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAAACACCT
351 CAACAAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGGC TACGAAGGCG
401 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGTAACCTAC

```



```

451 ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA
501 CGACAGCGAA ATGAAGCTCG TCGCTACAA CCAATCCGGC GCCGACTTCT
551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
601 GCGGTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT
651 ACCCGGCAGA GAAGATGCCT ACGGCAACCG TCCTTTCATC GCACAGGACG
701 ACCAAAATGC CCCCCTGTT CCGGCTGCGC GCCTCGGCTT CCACCTGAAA
751 GCCTCGCTGA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGT
801 CGCCCCAAAC AAACCTGCCC GCTACGAAAC GCGCACGCCC GGACACCATA
851 TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
951 CAGCAGCTTT CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACCGGCG
1001 GCGTGAACGT GAAGTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 598; ORF 149>:

```

m149.pep
  1 MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIQY DKALIDRENY
 51 YNHPLPDLGA HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFNGY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVSGDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQONAPRV PAARLGPHLK
251 ASLTDRIAN LDYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 149 shows 95.9% identity over a 339 aa overlap with a predicted ORF (ORF 149.ng) from *N. gonorrhoeae*:

m149/g149

	10	20	30	40	50	60
m149.pep	MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIQY DKALIDRENY YNHPLPDLGA					
g149	MLIDNNVRHYS SFFGVEQANW DNFTLEGGVR VEKQKASIRYDKALIDRENY NQPLPDLGA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m149.pep	HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA HGKHVATNTF EVGNKHLNKE					
g149	HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA HGKHVATNTF EVGNKHLNKE					
	70	80	90	100	110	120
	130	140	150	160	170	180
m149.pep	RSNNIELALG YEGDRWQYNL ALYRNRFNGY IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG					
g149	RSNNIELALG YKGRWQYNL AAYRNRFNGY IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG					
	130	140	150	160	170	180
	190	200	210	220	230	240
m149.pep	ADFYGAEGEI YFKPTPRYRIGVSGDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQONAPRV					
g149	ADFYGAEGEI YFKPTPRYRIGVSGDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQONAPRI					
	190	200	210	220	230	240
	250	260	270	280	290	300
m149.pep	PAARLGPHLK ASLTDRIAN LDYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW					
g149	PAARLGPHLK ASLTDRIAN LDYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW					
	250	260	270	280	290	300
	310	320	330	340		
m149.pep	NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKFX					

g149 NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFAGGVNVKFX  
310 320 330 340

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 599>:

a149.seq  
1 ATGCTGCTTG ACAATAAAGT GCAACATTAC AGCTTTTTTCG GTGTAGAACA  
51 GGCAAACTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAACTAC  
101 AAAAAAGCCTC CATCCGCTAC GACAAAGCAT TGATTGATCG GGAAAACTAC  
151 TACAACCATC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCCGCTC  
201 ATTCGCACTT TCGGGCAACT GGTATTTCAC GCCACAACAC AAACCTCAGCC  
251 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCAAGA GCTGTACGCA  
301 CACGGCAAAC ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAACACCT  
351 CAACAAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGGC TACGAAGCGC  
401 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGCAACTAC  
451 ATTTACGCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA  
501 CGACAGCGAA ATGAAGCTCG TCGCTACAA CCAATCCGGT GCGGACTTCT  
551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC  
601 GCGTTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT  
651 ACCCGGCAGG GAAGACGCCT ACGGCAACCG CCCACTCATT GCCCAAGCCG  
701 ACCAAAACGC CCCTCGCGTT CCGGCTGCGC GCCTCGGCGT CCACCTGAAA  
751 GCCTCGCTGA CCGACCGCAT CGATGCCAAT TTGGACTACT ACCGCGTGTT  
801 CGCCCAAAC AAACCTCGCC GCTACGAAAC GCGCACGCCG GGACACCATA  
851 TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG  
901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA  
951 CAGCAGCTTC CTCTCTGATA CGCCGCAAAAT GGGCCGCAGC TTTACGGGCG  
1001 GCGTGAACGT GAAGTTTTAA

This corresponds to the amino acid sequence <SEQ ID 600; ORF 149.a>:

a149.pep  
1 MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKOKASIRY DKALIDRENY  
51 YNHPLPDLGA HRQTARSFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA  
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFNGY  
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI  
201 GVSVDYVRGR LKNLPSLPGR EDAYGNRPLI AQADQNAPRV PAARLGVHLK  
251 ASLTDRIAN LDYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW  
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF\*

m149/a149 98.8% identity in 339 aa overlap

m149.pep	10	20	30	40	50	60
	MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKOKASIRY DKALIDRENY YNHPLPDLGA					
a149	MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKOKASIRY DKALIDRENY YNHPLPDLGA					
	10	20	30	40	50	60
m149.pep	70	80	90	100	110	120
	HRQTARSFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA HGKHVATNTF EVGNKHLNKE					
a149	HRQTARSFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA HGKHVATNTF EVGNKHLNKE					
	70	80	90	100	110	120
m149.pep	130	140	150	160	170	180
	RSNNIELALGY EGDWRWQYNLALYRNRFNGY IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG					
a149	RSNNIELALGY EGDWRWQYNLALYRNRFNGY IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG					
	130	140	150	160	170	180
m149.pep	190	200	210	220	230	240
	ADFYGAEGEI YFKPTPRYRIGVSGDYVRGR LKNLPSLPGR EDAYGNRPLIAQADQNAPRV					
a149	ADFYGAEGEI YFKPTPRYRIGVSGDYVRGR LKNLPSLPGR EDAYGNRPLIAQADQNAPRV					
	190	200	210	220	230	240
m149.pep	250	260	270	280	290	300
	PAARLGFHLKASLTDRIAN LDYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW					

**q149-1.seq**

1	ATGGCACAAA	TCACACTCAA	ACCATTGTGT	TTATCAATTC	TTTTAATCAA
51	CACACCCCTC	CTCGCCCAAG	CGCATGAAC	TGAGCAATCG	GTGGGCTTGG
101	AAACGGTCAG	CTGCTCGGC	AAAAGCGTC	CGCGCGCGAC	TTCCGGGCTG
151	CTGCACACTT	CGACCGCTC	GCACAAAATC	ATCTCCGGCG	ATACTTTGCG
201	CCAAAAAGCC	GTCAACTTGG	GGCAGCCTTT	GGACGGCGTA	CCGGGCATCC
251	ACGCTTCGCA	ATACGGCTGG	GGCGCATCCG	CTCCCGTTAT	TCGGCGCTCAA
301	ACGGGCAGAC	GATTAAAGT	ATTGAACCAT	CACGGCGAAA	CGGGCGATAT
351	GGCGGACTTT	TCTCCCAGTC	ACGCCATTAT	GGTAGATACC	GCCTTGTCG
401	AACAGGTTGA	AATCTGCGCG	GGGCGGCTTA	CGCTCTTTGA	CAGCTCGGGC
451	AATGTGGCGG	GCTGTGCTCA	TGTTCCGATG	GAAAAAATCC	CCGAAAAAAT
501	GCCTGAAAAC	GGCGTATCGG	GCGAagccgG	ATTGCGTTTG	ACGACGGCGCA
551	ATTTTAAAAA	ACTGACATCC	GCAGGCATCA	ATATCGGACT	GGGCAAAAAAC
601	TCGTGCTGTC	ATCCGAAGG	CTTGACCCG	AAATCGGGCG	ATTACCGCGT
651	ACCGCGTTAC	GCCAAATCTGA	AACGCCTGCC	GCACAGCGAT	CGCGATTTCG
701	AAACGGGCGAC	CATCGGGCTG	TCTTGGGTGG	GCGAAAAAGG	CTTTATCGGC
751	GCAGCATACA	GCGACCGTCG	CGACCGCTAC	GGCCTGCGCTG	CCCCACGCCA
801	CGAATCGCAT	GATTGCCACG	CGGACATTAT	CTGGGAAAAAG	AGTTTGATCA
851	ACAAACGCTA	TTTGACGTTT	TATCCGCATC	TGTTGACCGA	AGAAGACATC
901	GATTACGACA	ATCCGGGCTT	GAGCTCGCGG	TTCCACGACG	CGCGAGGTGC
951	ACACGCGACAC	ACCCACAACG	GCAAAACGTG	GATAGACCTG	GCGACCAAAAC
1001	GCTACGAACT	CCGCGCCGAA	TGGAAGCAGC	CATTCCCCGG	TTTTGAAGCC
1051	CTGCGCGTAC	ATCTGAACCG	CAATGACTAC	CACCACGACG	AAAAAGCAGG
1101	CGATGCGAGTA	GAAAACCTTC	TCAACAACAA	AACACCAAC	CGCCGTATCG
1151	AGTTGCGCCA	CCAACCCATA	GGCCGTCTGA	AAGGCAGATG	GGGCGTGCAA
1201	TATTTGGGAC	AAAAATCCAG	CGCGCTTTCG	GCCATTCCCCG	AAACCGTCCA
1251	ACAACCGATG	TTGATTGACA	ACAATTGTCG	CCATTACAGC	TTTTTCGGTG
1301	TAGAACAGGC	AAATTGGGAC	AACTTACAGC	TTGAAGCGCG	CTGACGGCTG
1351	GAAAAACAAA	AAGCCTCCAT	CCGGTACGAC	AAAGCATGCA	TGTATCGAGA
1401	AAACTACTAC	AACACGCCCC	TGCCCGACCT	CGGCGCTGCAC	CGCCAAACCC
1451	CCCGCTCGTT	CGCATTTTCG	GGCAACTTGT	ATTTACAGCC	ACACCAACAA
1501	CTCAGCCTCA	CCGCCCTCCA	TCAGGAACCG	CTGCCGTCAA	CGCAAGAACT
1551	GTACGCTACG	GGCAAGCAGC	TCGCCACCAA	CACCTTTGAA	GTCGGCAACA
1601	AACACCTCAA	CAAGAGCGCT	TCCAACAATA	TCGAACTCGC	GTCGGGCTAC
1651	GAAGGGCGAC	GCTGGCAATA	CAATCTGGCA	GCCTACCGCA	ACCGATTTCG
1701	CAACTACATT	TACGCCCAA	CCTTAAACGA	CGGACGGCGC	CCCCAATCCA
1751	TCGAGACAGCA	CAGCGAAATG	AACTCTGTGC	GCTACAACCA	ATCCGTTGCC
1801	GACTTCTACG	GCGCGGAAGG	CGAAATCTAC	TTCAAACCGA	CACCGCGCTA
1851	CCCGATCTGG	GTTCCTGGCG	ACTATTGACG	AGGCCGTCTG	AAAAAAGCTG
1901	CGTCCCTACC	CGGCAGGGAA	GATCCCTACG	GCAAAACGTG	CTTCATCGCA
1951	CAAGCGGACC	AAAACGCCCC	CCGATTTCGG	GCTGCGCGCC	TCGGCTTCCA
2001	CCTGAAAAAC	TCGCTAACCG	ACCGTATCGA	TGCCAATTTG	GACTACTACC
2051	CGGTGTTTCG	CCAAACAACA	CTCGCCCGCT	ACGAAACGCG	TACGCCCGGA
2101	CACCATATGC	TCAACCTCGG	TGCAAACTAC	CGCCGCAATA	CGCGCTATGG
2151	CGAGTGGAAAT	TGGTAGGTCA	AAGCGACAAA	CCTGCTCAAC	CAATCCGTTT
2201	ACGCCCAACG	CAGCTTCCTC	TCTGATACGC	CGCAATGGG	CCCGAGCTTT
2251	ACCGGCGGGC	TAAACGTGAA	GTTTTTAA		

**g149-1.pep**

1	MAQITLKPIV	LSILLINTPL	LAQAHETEQS	VGLETVSVVG	KSRPRATSGL
51	LHTSTASDKI	ISGDTLRQKA	VNLGDALDVG	PGIHASQOYGG	GASAPVIRRGQ
101	TGRRIKVLNH	HGETGDMADF	SPDHAIMDVT	ALSQQVEILR	GPVTLVLYSG
151	NVAGLVADV	GKIPEKMPEN	GVSGEAGLRL	SSGNLEKLTs	AGINIGLGKN
201	FVLHTEGLRY	KSGDVAVPRY	RNLKRLPDSH	ADsQTGSIGL	SWVGEGKFGI
251	AYSDNRDRRY	GLPAHSHEYD	DCHADI IWQK	SILNKRYLRL	YPHLLTEEDI
301	DYDNPGLSCG	FHDGDGAHAH	THNGKPWIDL	RNKRYELRAE	WOKFPFGEFA

```

351 LRVHLNRNDY HHDEKAGDAV ENFFNNKTHN ARIELRHQPI GRLKGSWGVQ
401 YLGQKSSALS AIPETVQQPM LIDNNVRHYS FFGVEQANWD NPTLEGGVRV
451 EKQKASIRYD KALIDRENYN NOPLPDLGAH RQTARSFALS GNWYFTPHHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY
551 EGDWRQYNLA AYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
601 DFYGAEGEII FKPTPRYRIG VSGDYVRGRL KNLPSPGPRE DPYGRPFPIA
651 QADQNAPRIP AARLGFHLKT SLTDRIDANL DYYRVFAQNK LARYETRTPG
701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
751 TGGVNVKF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 603>:

m149-1.seq

```

1 ATGGCACAAA CTACACTCAA ACCCATTGTT TTATCAATTC TTTAATCAA
51 CACACCCCTC CTCGCCAAG CGCATGAAAC TGAGCAATCG GTGGATTGCG
101 AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGGCCAC GTCGGGGCTG
151 TTGCACACTT CGACCGCCTC CGACAAAATC ATCTCCGGCG ATACCTTGCG
201 CCAAAAAGCC GTCACCTGG GCGACGCTTT AGACGGCGTA CCGGGCATCC
251 ACGCTTCGCA ATACGGCGGC GCGCGTCTG CTCCGTCAT TCGCGTCAA
301 ACAGGCAGGC GGATTAAAGT GTTGAACCAT CACGGCGAAA CAGGCGATAT
351 GGCGGATTTT TCGCCGATC ACGCCATTAT GGATAGATAC GCCTTGTCGC
401 AACAGGTCGA AATCCTGCGC GGGCCGGTTA CGCTCTTGTA CAGCTCGGGC
451 AATGTGGCGG GGCTGGTGA TGTGCGCAT GGCAAAATCC CCGAAAAAAT
501 GCCTGAAAAC GCGGTATCGG GCGAACTCGG ATTGCGTTTG AGCAGCGGCA
551 ATCTGGAAAA ACTCACGTCC GCGGCGATCA ATATCGGTTT GGGCAAAAAC
601 TTTGTATTGC ACACGGAAGG GCTGTACCGC AAATCGGGGG ATTACGCCGT
651 ACCGCGTTAC CGCAATCTGA AACGCCTGCC CGACAGCCAC GCCGATTTCG
701 AAACGGGCGC CATCGGGCTG TCTTGGGTTG GCGAAAAAGG TTTTATCGGC
751 GTAGCGTACA GCGACCGTCG CGACCAATAT GGTCTGCCTG CCCACAGCCA
801 CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGCTTGATTA
851 ACAAACGCTA TTTACAGCTT TATCCGCACC TGTTGACCGA AGAAGACATC
901 GATTACGACA ATCCGGGCTT GAGCTGCGGC TTCCACGACG ACGATAATGC
951 ACACGCACAC ACCACAGCG GCAGACCGTG GATAGACCTG CGCAACAAAC
1001 GCTACGAAT CCGTGCCGAA TGGAAGCAAC CGTCCCCCG TTTGAAGCC
1051 CTGCGCGTAC ACCTGAACCG CAACGACTAC CGCCACGACG AAAAAGCAGG
1101 CGATGCAGTC GAAAACCTTT TTAACAACCA AACGCAAAAC GCCCGCATCG
1151 AGTTGCGCCA CCAACCCATA GGTCTGTGTA AAGGCAGCTG GGGCGTGCAA
1201 TATTTACAAC AAAATCCAG TGCTTTATCT GCCATATCCG AAGCGGTTAA
1251 ACAACCGATG CTGCTTGACA ACAAGTGACA ACATTACAGC TTTTTCGGTG
1301 TAGAACAGGC AAACCTGGAC AACTTCACGC TTGAAGGAGG CGTACGCGTG
1351 GAAAAACAAA AAGCCTCCAT TCAGTACGAC AAAGCATTGA TTGATCGGGA
1401 AAACCTACTAC AACCAACCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1451 CCCGCTCATT CGCACTTTCC GGCAACTGGT ATTTACAGCC ACAACACAAA
1501 CTCAGCCTGA CCGCCTCCCA TCAGGAACGC CTGCCGTCAA CGCAAGAGCT
1551 GTACGCACAC GGCAACACG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1601 AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
1651 GAAGGCGACC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1701 TAACTACATT TACGCCAAA CCTTAAACGA CGGACGCGGC CCCAAATCCA
1751 TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGCGCC
1801 GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1851 CCGCATCGGC GTTTCGGCG ACTATGTACG AGGCCGCTG AAAAACCTGC
1901 CTTCCCTACC CGGCAGAGAA GATGCTTACG GCAACCGTCC TTTATCGCA
1951 CAGGACGACC AAAATGCCCC CCGTGTTCCG GCTGCGCGCC TCGGCTTCCA
2001 CCTGAAAGCC TCGCTGACCG ACCGTATCGA TGCCAATTG GACTACTACC
2051 CCGTGTTTCG CCAAAACAAA CTCGCCCGCT ACGAAACGCG CACGCCCGGA
2101 CACCATATGC TCAACCTCGG CGCAAACTAC CGCCGCAATA CGCGCTATGG
2151 CGAGTGGAAAT TGTACGTCA AAGCCGACAA CCTGCTCAAC CAATCCGTTT
2201 ACGCCACAG CAGCTTTCTC TGTATACG CGCAATGGG CCGCAGCTTT
2251 ACCGCGCGCG TGAACGTGAA GTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 604; ORF 149-1>:

m149-1.pep

```

1 MAQTTLKPIV LSILLINTPL LAQAHETEQS VDLETVSUVG KSRPRATSGI
51 LHTSTASDKI ISGDTLRQKA VNLGDALDGV PGIHASQYGG GASAPVIRGQ
101 TGRRIKVLNH HGETGDMADF SPDHAIMVDT ALSQQVEILR GPVTLLYSSG
151 NVAGLVDVAD GKIPKMPEN GVSGLGLRL SSGNLEKLTG GGINIGLGN
201 FVLHTEGLYR KSGDYAVPRY RNLKRLPDSH ADSQTGSIGL SWVGEKGFIG
251 VAYSRRDQY GLPAHSHEYD DCHADIIWQK SLINKRYLQL YPHLLTEEDI
301 DYDNPLGLSCG FHDDDNAHAH THSGRPWIDL RNKRYELRAE WKQFPFGFEA
351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTON ARIELRHQPI GRLKGSWGVQ
401 YLQKSSALS AISEAVKQPM LLDNKVQHYS FFGVEQANWD NPTLEGGVRV
451 EKQKASIQYD KALIDRENYN NHPLPDLGAH RQTARSFALS GNWYFTPOHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY

```

551 EGDWRQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA  
601 DFYGAEGEIIY FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPFFIA  
651 QDDQNAPRVP AARLGFLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG  
701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF  
751 TGGVNVKF\*

m149-1/g149-1 96.2% identity in 758 aa overlap

m149-1.pep	10	20	30	40	50	60
	MAQTTLKPIVLSILLINTPLLAQAHE	TEQSVDLET	VS	VVGKSRPRATS	G	LLHTSTASDKI
g149-1	MAQITLKPIVLSILLINTPLLAQAHE	TEQS	VGLETVS	VVGKSRPRATS	G	LLHTSTASDKI
	10	20	30	40	50	60
m149-1.pep	70	80	90	100	110	120
	ISGDTLRQKAVNLGDALDGVPGI	HASQYGGGASAPVIRG	QTGRRIKVLNHH	GETGDMADF		
g149-1	ISGDTLRQKAVNLGDALDGVPGI	HASQYGGGASAPVIRG	QTGRRIKVLNHH	GETGDMADF		
	70	80	90	100	110	120
m149-1.pep	130	140	150	160	170	180
	SPDHAIMVDTALSQQVEILRGPV	TLLYSSGNVAGLVDVADGKI	PEKMPENGVS	GELGLRL		
g149-1	SPDHAIMVDTALSQQVEILRGPV	TLLYSSGNVAGLVDVADGKI	PEKMPENGVS	GELGLRL		
	130	140	150	160	170	180
m149-1.pep	190	200	210	220	230	240
	SSGNLEKLTSGGINIGLGKNFV	LHTEGLYRKS	GDYAVPRYRNLKRLP	DSHAD	SQTGSIGL	
g149-1	SSGNLEKLTSAGINIGLGKNFV	LHTEGLYRKS	GDYAVPRYRNLKRLP	DSHAD	SQTGSIGL	
	190	200	210	220	230	240
m149-1.pep	250	260	270	280	290	300
	SWVGEKGFIVAYS	DRRDQYGLPAHSHEYDDCHADII	WQKSLINKRYLQ	LYPHLLTEEDI		
g149-1	SWVGEKGFIVAYS	DRRDQYGLPAHSHEYDDCHADII	WQKSLINKRYLQ	LYPHLLTEEDI		
	250	260	270	280	290	300
m149-1.pep	310	320	330	340	350	360
	DYDNPGLSCGFHDDNAHA	HTHSGRPWIDLRNKRYELRAEWKQ	FFPGFEALRVHLNRNDY			
g149-1	DYDNPGLSCGFHDDNAHA	HTHSGRPWIDLRNKRYELRAEWKQ	FFPGFEALRVHLNRNDY			
	310	320	330	340	350	360
m149-1.pep	370	380	390	400	410	420
	RHDEKAGDAVENFFNNQ	TQNARIELRHQPIGRLKGS	WGVQYLQKSSALSAISEAVKQPM			
g149-1	RHDEKAGDAVENFFNNQ	TQNARIELRHQPIGRLKGS	WGVQYLQKSSALSAISEAVKQPM			
	370	380	390	400	410	420
m149-1.pep	430	440	450	460	470	480
	LLDNKVQHY	SFFGVEQANWDNFTLEGGV	RVEKQKASIQYDKALIDRE	NYNHPDLGAH		
g149-1	LIDNNVRHY	SFFGVEQANWDNFTLEGGV	RVEKQKASIRYDKALIDRE	NYNQPLDLGAH		
	430	440	450	460	470	480
m149-1.pep	490	500	510	520	530	540
	RQTARSFALSGN	WYFTPHKLSLTASHQERLP	STQELYAHGKHVATNTFEVGNKHLN	KER		
g149-1	RQTARSFALSGN	WYFTPHKLSLTASHQERLP	STQELYAHGKHVATNTFEVGNKHLN	KER		
	490	500	510	520	530	540
m149-1.pep	550	560	570	580	590	600
	SNNIELALGYEGDRWQYNL	ALYRNRFGNYIYAQTLNDGRG	PKSIEDDSEM	KLVRYNQSGA		
g149-1	SNNIELALGYEGDRWQYNL	ALYRNRFGNYIYAQTLNDGRG	PKSIEDDSEM	KLVRYNQSGA		
	550	560	570	580	590	600
m149-1.pep	610	620	630	640	650	660
	DFYGAEGEIIYFKPTPRYRIG	VSGDYVRGRLKNLPSLP	GREDAYGNRPFFIAQDDQNAPRVP			
g149-1	DFYGAEGEIIYFKPTPRYRIG	VSGDYVRGRLKNLPSLP	GREDPYGRPFIAQADQNAPRIP			
	610	620	630	640	650	660

	670	680	690	700	710	720
m149-1.pep	AARLGFHLKASLTDRIDANLDYYRVFAQNKLRARYETRTPGHHMLNLGANYRRNTRYGEWN					
	:					
g149-1	AARLGFHLKTSLTDRIDANLDYYRVFAQNKLRARYETRTPGHHMLNLGANYRRNTRYGEWN					
	670	680	690	700	710	720
	730	740	750	759		
m149-1.pep	WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKF	X				
g149-1	WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKF	X				
	730	740	750			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 605>:

al49-1. seq

1	ATGGCACAAA	CTACACTCAA	ACCCATTGTT	TTATCAATTC	TTTAAATCAA
51	CACACCCCTC	CTCTCCCAAG	CGCATGGAAC	TGAGCAATCA	GTGGGCTTGG
101	AAACCGGTGAG	CGTCGTGCGG	AAAAGCGCTC	CGCGCGCCAC	TTCCGGGGCTG
151	CTGCACACTT	CTACCGCCTC	CGACAAAATC	ATCAGCGCGG	ACACCTTGGC
201	ACAAAAGAGC	GTCAACTTGG	GTGATGCTTT	AGACGGCGTA	CCGGGCATTG
251	ATGCCTCGCA	ATACGGCGGG	GGGCGATCCG	CTCCCGTTAT	TCGCGGTCAA
301	ACAGGCAGAC	GGATTAAAGT	TTGGAACCAT	CACGGCGGAA	CGGCGGACAC
351	GGCGGACTTC	TCTCCAGACC	ATGCAATCAT	GGTGACAGC	GCCTTGTCCG
401	AACAGGTGCA	AATCTCTGCG	GGTCCGGTTA	CGCTCTTGTA	CAGCTCGGGC
451	AATGTGGCGG	GGCTTGTGCA	TGTTGCGCAT	GGCAAAATCC	CCGAAAAAAT
501	GCCTGAAAGC	GGCGTATCGG	GCGAACTCGG	ATTGCGTTTG	AGCAGCGGCA
551	ATCTGGA AAA	ACTCACGTCG	GGCGGCGATC	ATATCGGTTT	GGGCAAAAAC
601	TTTGTATTGC	ACACGGGAAG	GCTGTACCGC	AAATCGGGGG	ATTACGCCGT
651	ACCGCGTTAC	CGCAATCTGA	AACGCCTGCC	CGACAGCCAC	CGCGATTTCG
701	AAACGGGGCG	CATCGGGGCT	TCTTGGGTTG	GCGAAAAAGG	CTTTATCGGC
751	GCAGCATACA	GCAGCGGCTG	CGACCAATAT	GGTCTGCCTG	CCCAACGCCA
801	CGAATACGAT	GATTGCCACG	CCGACATCAT	CTGGCAAAA	AGTTTGATTA
851	ACAAACCGTA	TTTGACAGTT	TATCCGCACC	TGTTGACCGA	AGAAGACATC
901	GATTACGACA	ATCCGGGCTT	GAGCTGCGGG	TTTACAGACG	ACGATGATGC
951	ACACGCCCCAT	GCCCCAACAG	GCAAACTCTG	GATAGACCTG	CGCAACAAAC
1001	GCTACGAACT	CCGCGCCGAA	TGGAAGCAAC	CGTTCCCGGG	TTTTGAAGCC
1051	GTCCGCGTAC	ACCTGAACCG	CAACGACTAC	CGCCACGACG	AAAAAGCATG
1101	CGATGCAGTA	GA AA ACTTTT	ATAACAACCA	ACCGCAAAAC	GCCCGTATCG
1151	AGTTGCGCCA	CCAACCATA	GGCCGTCTGA	AAGGCGACGT	GGGCGTGCAA
1201	TATTTGGGAC	AAA AATCCAG	TGCTTTATCT	GCCACATCCG	AAGCGGTCAA
1251	ACAACCGATG	CTGCTTGACA	ATAAAGTGCA	ACATTACAGC	TTTTTCGGTG
1301	TAGAACAGGC	AAACTGGGAC	AACTTACAGC	TTGAAGCGGG	CTGACGGGTG
1351	GAAAAACAAA	AAGCCTCCAT	CCGCTACGAC	AAAGCATTGA	TTGATCGGGA
1401	AAACTACTAC	AACCATCCCC	TGCCCGACCT	GGCGCGCGAC	CGCCAAACCG
1451	CCCGCTCATT	CGCACTTTTC	GGCAACTGGT	ATTTACAGCC	ACAAACAAA
1501	CTCAGCCTGA	CCGCCTCCCA	TCAGGAACGC	CTGCCGTCAA	CGCAAGAGCT
1551	GTACGCACAC	GGCAAAACGC	TCGCCACCAA	CACTTTGACA	GTCGGCAACA
1601	AACACCTCAA	CAAAGGACGT	TCCAACAATA	TCGAATTCGC	GCTGGGCTAC
1651	GAAGGCGACC	GCTGGCAATA	CAATCTGGCA	CTCTACCGCA	ACCGGTTTCG
1701	CAACTACATT	TACGCCCAAA	CCTTAAACGA	CGGACGGCGC	CCCAATTCGA
1751	TCGAAGACGA	CAGCGAAATG	AAGTCTGTGC	GTCACAACCA	ATCCGGTGGC
1801	GACTTCTACG	GCGGGAAGG	CGAAATCTAC	TTCAAAACCGA	CACCGCGCTA
1851	CCGCATCGGC	TTTCCGGGCG	ACTATGTACG	AGGCGGCTGT	AAAAACCTGC
1901	CTTCCCTACC	CGGCAAGGGA	GACGCTTACG	GCAACCGCCC	ACTCATTTGC
1951	CAAGCCGACC	AAAACGCCCC	TCGGTTCCTG	GCTGCGCGCC	TCGGCGTCCA
2001	CCTGAAGAGC	TCGCTGACCG	ACCGCATCGA	TGCCAATTTG	GACTACTACC
2051	GCGTGTTCGC	CCAAAACAAA	CTGCCCGCTG	ACGAAACCGC	CACGCCCCGA
2101	CACCATATGC	TCAACCTCGG	CGCAAACTAC	CGCCGCAATA	CGCGCTATGG
2151	CGAGTGGAA	TGCTACGTCTA	AAGCGACAAA	CCTGCTCAAC	CAATCCGTTT
2201	ACGCCACAG	CAGCTTCTCT	TCTGATACGC	CGCAAAATGG	CCGCAAGCTT
2251	ACCGGCGGCG	TGAACGTGAA	GTTTTAA		

This corresponds to the amino acid sequence <SEQ ID 606; ORF 149-1.a>:  
a149-1.pep

a149-1.pep

1	MAQTTLKPIV	LSILLINTPL	LSQAHGTEQS	VGLETVSVVG	KSRPRATSGL
51	LHTSTASDKI	IGSDTLRQKA	VNLGDALDGV	PGIHASQYGG	GASAPVIRGQ
101	LGRRIKLVNH	HGETGMDADF	SPDHAIMVDS	ALSQQVEILR	GPVTLLYSSG
151	NVAGLVDDAD	GKIPEKMPEN	GVSSELGLRL	SSGNLEKITS	GGINIGLGNK
201	VLHTEGLYR	KSGDYAVPRY	RNLKRLPDSH	ADSQTGSIGL	SWVGEKGFIG
251	AAYSRRDRQY	GLPAHSHEYD	DCHADIIWQK	SILNKRYLQL	YPHLLTTEEDI
301	DYDNPGLSCG	FHDDDDAHAA	AHNGKPWIDL	RNKRYELRAE	WKQSPFFGEA

351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTON ARIELRHQPI GRLKGSWGVO  
401 YLGQKSSALS ATSEAVKQPM LLDNKVQHYS FFGVEQANWD NFTLEGGVRV  
451 EKQKASIRYD KALIDRENNY NHPLPDLAGH RQTARSAFALS GNWYFTPQHK  
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY  
551 EGDWQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA  
601 DFYGAEGEYI FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPLIA  
651 QADQNAPRVP AARLGVHLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG  
701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF  
751 TGGVNVKF\*

a149-1/m149-1 98.0% identity in 758 aa overlap

	10	20	30	40	50	60
a149-1.pep	MAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLETVSVVGKSRPRATSGLLHTSTASDKI					
m149-1	MAQTTLKPIVLSILLINTPLLAQAHETEQSVLDLETVSVVGKSRPRATSGLLHTSTASDKI					
	10	20	30	40	50	60
	70	80	90	100	110	120
a149-1.pep	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
m149-1	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
	70	80	90	100	110	120
	130	140	150	160	170	180
a149-1.pep	SPDHAIMVDSALSQQVEILRGPVTLTYSSGNVAGLVVDVADGKIPEKMPENGVSSELGLRL					
m149-1	SPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVVDVADGKIPEKMPENGVSSELGLRL					
	130	140	150	160	170	180
	190	200	210	220	230	240
a149-1.pep	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQTGSIGL					
m149-1	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQTGSIGL					
	190	200	210	220	230	240
	250	260	270	280	290	300
a149-1.pep	SWVGEKGFIGAAYSDDRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI					
m149-1	SWVGEKGFIVAYSDDRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI					
	250	260	270	280	290	300
	310	320	330	340	350	360
a149-1.pep	DYDNPGLSCGFHDDDDAHAAHNGKPWDLRNKRYELRAEWKQFPFGFEALRVHLNRNDY					
m149-1	DYDNPGLSCGFHDDDDAHAAHNGKPWDLRNKRYELRAEWKQFPFGFEALRVHLNRNDY					
	310	320	330	340	350	360
	370	380	390	400	410	420
a149-1.pep	RHDEKAGDAVENFFNNQTONARIELRHQPIGRLKGSWGVOYLGQKSSALSATSEAVKQPM					
m149-1	RHDEKAGDAVENFFNNQTONARIELRHQPIGRLKGSWGVOYLGQKSSALSATSEAVKQPM					
	370	380	390	400	410	420
	430	440	450	460	470	480
a149-1.pep	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENNYNHPLPDLAGH					
m149-1	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENNYNHPLPDLAGH					
	430	440	450	460	470	480
	490	500	510	520	530	540
a149-1.pep	RQTARSAFALSGNWFYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER					
m149-1	RQTARSAFALSGNWFYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER					
	490	500	510	520	530	540
	550	560	570	580	590	600
a149-1.pep	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSGA					
m149-1	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSGA					
	550	560	570	580	590	600
	610	620	630	640	650	660

```

a149-1.pep  DFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPLIAQADQNA PRVP
m149-1       DFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNA PRVP
              610      620      630      640      650      660

a149-1.pep  AARLGVHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEWN
m149-1       AARLGVHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEWN
              670      680      690      700      710      720

a149-1.pep  WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
m149-1       WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
              730      740      750      759

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 607>:

```

g150.seq (partial)
1  ..TACTGCAAGG CAGACCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAT
51  CACCGCCCGC CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTTGA
101 GCGGTTTCGGA TTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGGTTT
151 GACAACGATC CGGCACTGGT CGGGGAAATC CTAGACCTGC TCGGCATCAA
201 TCCGGCAACG GAAATACAGG CGGGCGGAAA AACCTGCGG GTTGCCCTCCG
251 CACTGTTATC CCATTTTCGAA CTCACGCAA ACACCCCGC CTTTGTCAAA
301 GGCTATGCCA CGTTCGCCGA TAATGACGAA CTCGACCGTA TTGCTGCCGA
351 CAACGCCGTT TTGCAAGGCT TTGTGCAAAG CACGCCGATT GCCCGTGTGC
401 TGCACCGCTT CCCGGCAAAA CTGACGGCGG AACAATTGCG CGGCCTGCTG
451 CGCCCGCTTG CGCCGCGCCT GTATTGCAAT TCCTCGTCGC AGGCGGAAGC
501 GGGGGACGAA GTGCACCTGA CCGTCGGCGC AGTGCCTTTC GAACACGAAG
551 GCGCGGCCAG GCGGGCGGCG GCATCGGGTT TCTTTGCCGA CCGGCTGGAA
601 GAGGACGGCA CGGTGCGCGT GTTTGCGGAA CGAACGACG GCTTCAGGCT
651 GCCCGAAGAC AGCCGCAAGC CGATTGTGAT GATCGGCTCC GGTACCGCG
701 TCGCACCGTT CCGCGCCTTC GTCCAACAAC GTGCCGAGA AAATGCGGAA
751 GGCAGAAACT GGCTGATTTT CCGCAATCCG CATTTGCCG CCGACTTCCT
801 CTATCAGACC GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGAT
851 ATGACTTCGC CTGGTCGCGC GATCAGGAAG AAAAAATCTA TGTGCAGGAC
901 AAAATCCGCG AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC
951 GCATATCTAT GTGTGCGGCG ATGCGGCAA AATGGCAAAA GAAGTGGAAG
1001 CCGCCTTGCT GGATGTGATT ATCGGGGCG GGCATTGCGA CGAAGACGGC
1051 GCAGAAGGAT ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA
1101 TGTTTATTGA

```

This corresponds to the amino acid sequence <SEQ ID 608; ORF 150.ng>:

```

g150.pep (partial)
1  ..YCKADFPFAA LLANQKITAR QSDKDVRHIE IDLSGSDLHY LPGDALGVWF
51  DNDPALVGEI LDLLGINPAT EIQAGGKTLV VASALLSHFE LTQNTPAFVK
101 GYATFADNDE LDRIAADNAV LQGFVQSTPI AGVLHRFPK LTAEQFAGLL
151 RPLAPRLYSI SSSQAEAGDE VHLLTVGAVRF EHEGRARAGG ASGFFADRLE
201 EDGTVRVFAE RNDGFRLPED SRKPIVMIGS GTGVAPFRAF VQQRAAENAE
251 GRNWLIFGNP HFAADFLYQT EWQQFAKDFG LHRYDFAWSR DQEEKIYVQD
301 KIREQAEGLW OWLQEGAHY VCGDAAKMAK EVEAALLDVI IGAGHSEDDG
351 AEGYLDMLRE EKRYQRDVY*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 609>:

```

m150.seq
1  ATGCAGAACAA CAAATCCGCC ATTACCGCCT CTGCCGCCCG AAATCACGCA
51  GCTCCTGTCG GGGCTGGACG CGGCACAAATG GCGTGGCTG TCCGGCTACG
101 CTTGGGCAAAA AGCAGGAAAC GGGGCATCTG CAGGACTGCC CGCGCTTCAG
151 ACGGCATTGC CGGCGGACGA ACCTTTTTCG GTAACCGTCC TTTCCGCTC
201 GCAAACCGGC AATGCGAAAT CCGTTGCCGA CAAAGCGGCG GACAGCCTGG
251 AAGCCGCCGG CATCCAAGTC AGTCGCGCCG AACTGAAAGA CTATAAGGCG
301 AAAAAATCATG CCGGCGAAGC CCGCCTGCTG CTGGTTACCT CCACCCAAGG
351 CGAAGGCGAA CCGCGGAAAG AAGCCGTCGT GCTGCACAAA CTGCTGAACG
401 GCAAAAAAGC CCCGAAATTG GACAAACTCC AATTGCGCGT ACTGGGTTTG

```



```

451  GCGGACAGTT  CCTATCCGAA  TTTCTGTCAG  GCAGGTAAAG  ATTTGACCGG
501  GCGTTTTGAA  GAATTGGGCG  CAAAACGGCT  GCTCGAACGC  GTTGATGCGG
551  ATTTGGACTT  TACCGCCTCC  GCAAACGCCT  GGACAGATAA  TATCGCCGCA
601  CTCTTAAAG  AAGAAGCCGC  AAAAAACCGG  GCAACGCCCC  CGCCGCAGAC
651  AACGCCCCC  GCCGGCCTTC  AGACGGCACC  GGATGGCAGG  TACTGCAAGG
701  CAGCCCCCTT  TCCCGCCGCC  CTGCTGGCCA  ATCAGAAAAT  CACCGCCCGC
751  CAATCCGATA  AAGACGTGCG  CCACATCGAA  ATCGATTTGA  GCGGTTCCGA
801  TTTGCACTAC  CTCCCGGGCG  ACGCGCTCGG  CGTTTGGTTT  GACAACGATC
851  TGGCACTGGT  CAGGGAAATC  CTAGACCTGC  TCGGCATCGA  TCCGGCAACG
901  GAAATACAGG  CGGGCGGAAA  GATGATGCCG  GTTGCGCGCG  CACTTTCATC
951  TCATTTGCAA  CTCACGCAA  AACTCCGGC  TTTCTGCAAA  GGCTATGCCG
1001 CGTTCGCCCC  TTATGAAGAA  CTCGATAAAA  TCATTGCCGA  TAACGCCGTT
1051 TTGCAGGATT  TCGTGCAAAA  CACGCCTATT  GTCGATGTGC  TGCACCGCTT
1101 CCCGGCAAGC  CTGACGGCAG  AACAATTCAT  CCGTTACTTG  CGTCCGCTTG
1151 CACCCCGTTT  GTATTGATT  TCTTCAGCAC  AGGCGGAAGT  GGGCGATGAA
1201 GTGCATTTAA  CTGTCGGCGT  GGTTCGTTTT  GAACACGAAG  GCCGCGCCAG
1251 AACGGCGGCG  GCATCGGGTT  TCCTTGCCGA  CCGGCTGGAA  GAGGACGGCA
1301 CGGTGCGCGT  GTTTGTGGAA  CGCAACGACG  GCTTCAGGCT  GCCGAAGAC
1351 AGCCGCAAGC  CGATTGTGAT  GATCGGCTCG  GGCACCGGCG  TCGCACCGTT
1401 CCGCGCTTTC  GTCCAACAAC  GTGCCGAGA  AAATGCGGAA  GGCAAAACT
1451 GGCTGATTTT  CGGCAATCCG  CATTTTGCCC  GTGATTTTCT  CTATCAAACC
1501 GAATGGCAGC  AGTTTGCCAA  AGACGGCTTC  CTGCACAGGT  ACGATTTTCG
1551 CTGGTCCCGC  GATCAGGAAG  AAAAAATCTA  TGTGCAGGAC  AAAATCCGCG
1601 AACAGGCGGA  AGGACTTTGG  CAATGGCTGC  AGGAAGGCGC  GCATATCTAT
1651 GTGTGCGGCG  ATGCGGCAAA  AATGGCAAAA  GACGTGGAAG  CCGCCTTGCT
1701 GGATGTGATT  ATCGGGGCG  GACATTTGGA  CGAAGAGGGC  GCAGAAGAAT
1751 ATTTGGATAT  GCTGCGCGAA  GAAAAACGCT  ATCAGCGTGA  TGTTTATTGA

```

This corresponds to the amino acid sequence <SEQ ID 610; ORF 150>:

m150.pep

```

1  MQNTNPPLPP  LPPEITQLLS  GLDAAQWAWL  SGYAWAKAGN  GASAGLPALO
51  TALPAAEPFS  VTVLSASQTG  NAKSVADKAA  DSLEAAGIQV  SRAELKDYKA
101 KNIAGERRLL  LVTSTQGEGE  PPKEAVVLHK  LLNGKKAPKL  DKLQFAVLGL
151 GDSSYPNFCQ  AGKDFDRRFE  ELGAKRLLER  VDADLDFTAS  ANAWTDNIAA
201 LLKEEAANKR  ATPAPQTPP  AGLQTAPDGR  YCKAAPFPAA  LLANQKITAR
251 QSDKDVRHIE  IDLSGSDLHY  LPGDALGVWF  DNDPALVREI  LDLLGIDPAT
301 EIQAGGKMMP  VARALSSHFE  LTQNTPAFVK  GYAAFAHYEE  LDKIIADNAV
351 LQDFVQNTPI  VDLVHRFPAS  LTAEQFIRLL  RPLAPRLYSI  SSAQAEVGDE
401 VHLTVGVVRF  EHEGRARTGG  ASGFLADRLE  EDGTVRVFVE  RNDGFRLPED
451 SRKPIVMIGS  GTGVAPFRAF  VQORAENAE  GKNWLIFGNP  HFARDFLYQT
501 EWQQFAKDF  LHYRDFAWSR  DQEEKIYVQD  KIREQAEGWL  QWLQEGAHYI
551 VCGDAAKMAK  DVEAALLDVI  IGAGHLDEEG  AEEYLDMLRE  EKRYQRDVY*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 150 shows 91.3% identity over a 369 aa overlap with a predicted ORF (ORF 150.ng) from *N. gonorrhoeae*:

m150/g150

```

                210      220      230      240      250      260
m150.pep  LLKEEAANKRATPAPQTPPAGLQTAPDGRYCKAAPFPAA LLANQKITARQSDKDVRHIE
                |||  |||  |||  |||  |||  |||
g150              YCKADPFPAALLANQKITARQSDKDVRHIE
                10      20      30

                270      280      290      300      310      320
m150.pep  IDLSGSDLHYLPGDALGVWFDNDPALVREILDLLGIDPATEIQAGGKMMPVARALSSHFE
                |||  |||  |||  |||  |||  |||
g150              IDLSGSDLHYLPGDALGVWFDNDPALVGEILDLLGINPATEIQAGGKTLFVASALLSHFE
                40      50      60      70      80      90

                330      340      350      360      370      380
m150.pep  LTQNTPAFVKGYAAFAHYEELDKIIADNAVLQDFVQNTPIVDVLRFPASLTAEQFIRLL

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g150  LTQNTPAFVKGYATFADNDELDRIAADNAVLQGFVQSTPIAGVLHRFPAKLTAEQFAGLL
      100      110      120      130      140      150
      390      400      410      420      430      440
m150.ppep RPLAPRLYSISSAQAEVGVDEVHVTGVVRFEGHEGRARTGGASGFLADRLEEDGTVRVFVE
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g150  RPLAPRLYSISSQAEAGDEVHVTGAVRFEHEGRARAGGASGFFADRLEEDGTVRVFAE
      160      170      180      190      200      210
      450      460      470      480      490      500
m150.ppep RNDGFRLPEDSRKPIVMIGSGTGAVPFRAFVQQRAAENAEGKNWLIFGNPHFARDFLYQT
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g150  RNDGFRLPEDSRKPIVMIGSGTGAVPFRAFVQQRAAENAEGKNWLIFGNPHFAADFLYQT
      220      230      240      250      260      270
      510      520      530      540      550      560
m150.ppep EWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGWLQEWLQEGAHYVCGDAAKMAK
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g150  EWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGWLQEWLQEGAHYVCGDAAKMAK
      280      290      300      310      320      330
      570      580      590      600
m150.ppep DVEAALLDVIIGAGHLDEEGAEYLDMLREEKRYQRDVYX
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g150  EVEAALLDVIIGAGHSDEEDGAEGYLDMLREEKRYQRDVYX
      340      350      360      370

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 611>:

```

a150.seq
1  ATGCAGAACA CAAATCCGCC ATTACCGCCT ATGCCGCCCG AAATCACGCA
51  GCTCCTGTCT GGGCTGGACG CGGCACAATG GGCGTGGCTG TCCGGCTACG
101 CTTGGGCAAA AGCAGGAAAC GGGGCATCTG CAGGACTGCC CGCGCTTCAG
151 ACGGCATTGC CGACGGCAGA ACCTTTTTCG GTAACCGTCC TTTCCGCCTC
201 GCAAACCGGC AATGCGAAAT CCGTTGCCGA CAAAGCGGCG GACAGCCTGG
251 AAGCCGCCCG CATCCAAGTC AGTCGCGCCG AACTGAAAGA CTATAAGGCG
301 AAAAACATCG CCGGCGAAGC CCGCCTGCTG CTGGTTACCT CCACCCAAGG
351 CGAAGGCGAA CCGCCGGAAG AAGCCGTCGT GCTGCACAAA CTGCTGAACG
401 GCAAAAAAGC CCCGAAATTG GACAACTCC AATTGCGGT ACTGGGTTTG
451 GGCGACAGCT CCTATCCGAA TTTCTGCCGG GCGGGCAAAG ATTTTCGACAA
501 ACGTTTTGAA GAATTGGGCG CAAAACGCCT GCTCGAACGC GTTGATGCGG
551 ATTTGGACTT TGCCGCCGCC GCAGACGGAT GGACAGATAA TATCGCCGCA
601 CTCTTAAAG AAGAAGCCGC AAAAACCCG GCAACGCCCG CGCCGCAGAC
651 AACGCCCCC GCGGCCTTC AGACGGCACC GGATGGCAGG TACTGCAAGG
701 CAGACCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAT CACCGCCCGC
751 CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTGTA GCGGTTCCGA
801 TTTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGTTT GACAACGATC
851 CGGCACTGGT CAGGGAATC CTAGACCTGC TCGGCATCGA TCAGGCAACG
901 GAAATACAGG CGGGCGGAAA AACCCTGCCG GTTGCTCCG CACTGTTATC
951 CCATTTTGAA CTCACGCAA ACACCCCGC CTTTGTCAA GGCTATGCCC
1001 CGTTCGCCGA TGATGACGAA CTCGACCGTA TTGCTGCCGA CAACGCCGTT
1051 TTGCAAGGCT TTGTGCAAAG CACGCCGATT GCCGATGTGC TGCACCGCTT
1101 CCCGGCAAAA CTGACAGCGG AACAATTGCG CGGCCTACTG CGCCCGCTTG
1151 CGCCGCGCCT GTATTCGATT TCCTCGTCGC AGGCGGAAGT GGGGGACGAA
1201 GTGCACTGA CCGTCGCGCG GGTGCGTTT GAACACGAAG GGCAGCCAG
1251 GCGGGGCGGC GCATCGGGT TCCTTGCCGA CCGGCTGGAA GAGGACGGCA
1301 CGGTGCGCGT GTTTGTGGA CGCAACGACG GCTTCAGGCT GCCCGAAGAC
1351 AGCCGCAAGC CGATTGTGAT GATCGGCTCG GGCACCGCG TCGCACCGTT
1401 CCGCGCTTTC GTCCAACAAC GTGCCGCGA AAATGCGGAA GGCAAAAAC
1451 GGCTGTTTTT CGCAATCCG CATTTGCCG GTGATTTTCT CTATCAAACC
1501 GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGAT ACGATTTCCG
1551 CTGGTTCGCG GATCAGGAAG AAAAATCTA TGTGCAGGAC AAAATCCGCG
1601 AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCG GCATATCTAT
1651 GTGTGCGGCG ATGCGGCAAA AATGGCAAAA GACGTGGAAG CCGCCTTGCT

```

1701 GGATGTGATT ATCGGGGCAG GACATTTGGA CGAAGAGGGC GCAGAAGAAT  
1751 ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTTTATTGA

This corresponds to the amino acid sequence <SEQ ID 612; ORF 150.a>:

a150.pep  
1 MQNTNPPLPP MPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ  
51 TALPTAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA  
101 KNIAGERRLL LVTSTQGEGE PPEEAVVLHK LLNGKKAPKL DKLQFAVLGL  
151 GDSSYPNFCR AGKDFDKRFE ELGAKRLLER VDADLDFAAA ADGWTDNIAA  
201 LLKEEAAKNR ATPAPQTPP AGLQAPDGR YCKADPFPA LLANQKITAR  
251 QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI LDLLGIDQAT  
301 EIQAGGKTL PVASALLSHFE LTQNTPAFVK GYAPFADDDE LDRIAADNAV  
351 LQGFVQSTPI ADVLHRFPK LTAEQFAGLL RPLAPRLYSI SSSQAEVGDE  
401 VHLTVGAVRF EHEGRARAGG ASGFLADRLE EDGTVRVFVE RNDGFRLPED  
451 SRKPIVMIGS GTGVAPFRAF VQQRAAENAE GKNWLFFGNP HFARDFLYQT  
501 EWQQFAKDG LHRDYFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHY  
551 VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY\*

m150/a150 94.8% identity in 599 aa overlap

	10	20	30	40	50	60
m150.pep	MQNTNPPLPPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQ	TALPTAEPFS	NAKSVADKAA	DSLEAAGIQV	SRAELKDYKA	KNIAGERRLL
a150	MQNTNPPLPPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQ	TALPTAEPFS	NAKSVADKAA	DSLEAAGIQV	SRAELKDYKA	KNIAGERRLL
	70	80	90	100	110	120
m150.pep	VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLL	LVTSTQGEGE				
a150	VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLL	LVTSTQGEGE				
	130	140	150	160	170	180
m150.pep	PPKEAVVLHKLLNGKKAPKLDKLQFAVLGLGDSSYPNFCQAGKDFDR	RFEELGAKRLLER				
a150	PPKEAVVLHKLLNGKKAPKLDKLQFAVLGLGDSSYPNFCQAGKDFDR	RFEELGAKRLLER				
	190	200	210	220	230	240
m150.pep	VDADLDFASANAWTDNIAALLKEEAAKNRATPAPQTPPAGLQAPDGR	YCKADPFPA				
a150	VDADLDFASANAWTDNIAALLKEEAAKNRATPAPQTPPAGLQAPDGR	YCKADPFPA				
	250	260	270	280	290	300
m150.pep	LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFDNDPALV	REILDLLGIDPAT				
a150	LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFDNDPALV	REILDLLGIDPAT				
	310	320	330	340	350	360
m150.pep	EIQAGGKMPVARALSSHFEELTQNTPAFVKGYAAFAHYEELDKIIAD	NAVLQDFVQNTPI				
a150	EIQAGGKMPVARALSSHFEELTQNTPAFVKGYAAFAHYEELDKIIAD	NAVLQDFVQNTPI				
	370	380	390	400	410	420
m150.pep	VDVLRFPASLTAEQFIRLLRPLAPRLYSISSAQAEVGDEVHDTVGV	RFEHEGRARTGG				
a150	VDVLRFPASLTAEQFIRLLRPLAPRLYSISSAQAEVGDEVHDTVGV	RFEHEGRARTGG				
	430	440	450	460	470	480
m150.pep	ASGFLADRLEEDGTVRVFVERNDGFRLPEDSRKPIVMIGSGTGVA	PFRAFVQQRAAENAE				
a150	ASGFLADRLEEDGTVRVFVERNDGFRLPEDSRKPIVMIGSGTGVA	PFRAFVQQRAAENAE				
	490	500	510	520	530	540
m150.pep	GKNWLIFGNPHFARDFLYQTEWQFAKDGFLHRYDFAWSRDQEEKIY	VQDKIREQAEGLW				

```

a150      GKNWLF FGNPHFARDFLYQTEWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGLW
           490      500      510      520      530      540
m150.pep  QWLQEGAH IYVCGDAAKMAKDVEAALLDVII GAGHLDEEGAE EYLDMLREEKRYQRDVYX
           550      560      570      580      590      600
a150      QWLQEGAH IYVCGDAAKMAKDVEAALLDVII GAGHLDEEGAE EYLDMLREEKRYQRDVYX
           550      560      570      580      590      600

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 613>:

```

g151.seq
1  ATGAAACAAA TCCGCAACAT CGCCATCATC GCACACGTCG ACCACGGCAA
51  AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGCGACA TTCCGCGCCA
101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTG
201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GCGGCGGAAG
251 TGGAGCGCGT TTTGGGGATG GTGGATTGCG TCGTCTTGTT GGTGGACGCA
301 CAGGAAGGTC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAACCGTCCG
401 CCCGTCCGAG CTGGGTTATC GACCAGACTT TCGAGTTGTT CGACAACCTG
451 GGTGCGACCG ACGAGCAGTT GGATTTCCTG ATTGTTTACG CTTCAGGTTT
501 GAGCGGCTTT GCCAAGCTGG AAGAAAccga CGAGAGCAGC GATATGCGCC
551 CGCtggttcga CACCATCCTA AAATACAcgc ctgCACCAG CCGCAGCGCG
601 GACGAGCCGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
651 CGGCCGCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
701 AAACCGTTGC CGTGATGAAC CACGAGCAGC AAATCGCCCA AGGCCGCATC
751 AACCAGCTTT TGGGTTTCAA AGGCTTGGA CCGGTGCCGC TTGAAGAAGC
801 CGAAGCCGCG GACATTGTGA TTATTTCCGG TATCGAAGAC ATCGGCATCG
851 GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCCTGCC GATGTTGAGC
901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTAAACA CCAGCCCCTG
951 CGCAGGTACA GAAGGCAAAT TCGTGACCAG CCGCCAAATC CGCGACCGCC
1001 TGCAAAAAGA ATTGCTGACC AACGTTGCCC TGCGCGTGGA AGACACCGCC
1051 GatgCCGACG TGTTCCGCGT ATCcgGGCGC GGCGAACGTC ACCTGACGAT
1101 TTTGTGGA AATATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAGC
1151 CGCGCGTCGT GTACCGAGAC ATCGACGCTC AAAAAATGCGA ACCTTATGAA
1201 AACCTGACTG TGGACGTACc cgacgacAAC CAAGGCGCGG TAATGGAAGA
1251 ACTCGGCCGC CGCCGTGGCG AACTGACCAA TATGGAAAGC GACGGCAACG
1301 GacgCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
1351 CAAGGCGAAT TCATGACCCT GACGCGCGGC GTCGGGCTGA TGAgccacGT
1401 GTTcgacgac tacgcgccg tcaAACCCGA TATGCCCGGC CGCCACAACG
1451 GCGTactggt GtcccaAGAG CAGGCGGAGG CGGTGCTTA CGCCTTGTGG
1501 AATCTTGAAG ACCGCGGCCG TATGTTTCGTA TCGCCCAACG ACAAAATCTA
1551 CGAAGGTATG ATTATCGGCA TCCACAGCCG CGACAACGAT TTGGTGGTCA
1601 ACCCGCTCAA AGGCAAAAAA CTCACCAATA TCCGTGCCAG CGGTACCGAC
1651 GAAGCGGTGC GCCTGACCAC GCCGATCAAA CTGAcgcTGG AAGGCGCGGT
1701 CGAGTTTATC GACGATGACG AGCTGGTGGA AATCAGCCG CAATccatcc
1751 gcctgcgcat gcgttacctG AGCGaattgg aacgcccgcg tcaTTTAA
1801 AagctgGATT AA

```

This corresponds to the amino acid sequence <SEQ ID 614; ORF 151.ng>:

```

g151.pep
1  MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
51  RGITILAKNT AIDYEGCHIN IVDTPGHADF GGEVERVLGM VDCVLLVDA
101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
151 GATDEQLDFP IVYASGLSGF AKLEETDESS DMRPLFDITL KYTPAPSGSA
201 DEPLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQTVAVMN HEQQIAQGR I
251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLOKELLT NVALRVEDTA
351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
401 NLTVDPDDN QGAVMEELGR RRGELTNMES DGNGRTRLEY HIPARGLIGF
451 QGEFMTLTRG VGLMSHVFD D YAPVKPDMPG RHNGVLVSQE QGEAVAYALW

```

501 NLEDGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD  
 551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRMYL SELERRRHF  
 601 KLD\*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 615>:

m151.seq

```

1  ATGAAACAAA TCCGCAACAT CGCCATCATC GCCACGTCG ACCACGGCAA
51  AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGGTA
201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG
251 TAGAGCGCGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG
301 CAGAAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGCG
351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAGCCGTCGG
401 CTCGTCCGAG CTGGGTATAT GACCAAACCT TCGAGCTGTT CGACAACTTG
451 GGCGCGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTACG CTTCAGGGTT
501 GAGCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
551 CGCTGTTTCA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
601 GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
651 CGGCGCGCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
701 AAACCGTTGC CGTCATGAAC CACGATCAGC AAATCGCCCA AGGCCGATC
751 AACCAGCTTT TGGGTTTCAA AGGTTTGGA CGCGTGCCGC TTGAAGAAGC
801 CGAAGCCGGC GACATCGTGA TTATTTCCGG TATCGAAGAC ATCGGTATCG
851 GCGTAACCAT CACCGACAAA GACAATCCCA AAGGCTACC GATGTTGAGC
901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGCT
951 GGCGGGTACG GAAGGCAAAT TCGTAACCAG CCGCAAATC CGCGACCGCC
1001 TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC
1051 GATGCCGACG TGTTCCGCGT ATCCGGGCGC GGCGAGCTGC ACCTGACCAT
1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC
1151 CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
1201 AACCTGACCG TGGATGTACC CGACGACAAC CAAGCGCGCG TAATGGAAGA
1251 ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
1301 GACGACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
1351 CAAGGCGAAT TTATGACCCT GACGCGCGGG GTCGGGTGTA TGAGCCAGCT
1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCCGGC CGCCACAACG
1451 GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTTG
1501 AATCTGGAAG ACCGCGGCCG TATGTTGTA TCGCCAACG ACAAATCTA
1551 CGAAGGCATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
1601 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
1651 GAAGCCGTTT GCCTGACCAC GCCAATCAAG CTGACGCTGG AAGGTGCGGT
1701 TGAGTTTATC GACGATGACG AACTCGTTGA AATCACGCCG CAATCCATCC
1751 GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCCGCG CCACTTTAAA
1801 AAGCTGGATT GA

```

This corresponds to the amino acid sequence <SEQ ID 616; ORF 151>:

m151.pep

```

1  MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
51  RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVLLVDA
101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDLIL KYTPAPSGSA
201 DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQTVAVMN HDQQIAQGRI
251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLOKELLT NVALRVEDTA
351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
401 NLTVDVDDN QGAVMEELGR RRGELTNMES DNGRTRLEY HIPARGLIGF
451 QGEFMTLTRG VGLMSHVDD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW
501 NLEDGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRKRYL SELERRRHF
601 KLD*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 151 shows 99.2% identity over a 603 aa overlap with a predicted ORF (ORF 151.ng) from *N. gonorrhoeae*:

m151/g151

m151.pep	10	20	30	40	50	60
	MKQIRNIAIIAHVDHGKTTLVLDQLLRQSGTFRANQQVDERVMSNDLEKERGITILAKNT					
g151	MKQIRNIAIIAHVDHGKTTLVLDQLLRQSGTFRANQQVDERVMSNDLEKERGITILAKNT					
	10	20	30	40	50	60
m151.pep	70	80	90	100	110	120
	AIDYEGYHINIVDTPGHADFGGEVERVLGMVDCVLLVDAQEGPMPQTRFVTKKALALGL					
g151	AIDYEGCHINIVDTPGHADFGGEVERVLGMVDCVLLVDAQEGPMPQTRFVTKKALALGL					
	70	80	90	100	110	120
m151.pep	130	140	150	160	170	180
	KPIVVINKIDKPSARPSWVIDQTFELFDNLGATDEQLDFPIVYASGLSGFAKLEETDESN					
g151	KPIVVINKIDKPSARPSWVIDQTFELFDNLGATDEQLDFPIVYASGLSGFAKLEETDESS					
	130	140	150	160	170	180
m151.pep	190	200	210	220	230	240
	DMRPLFDITLKYTPAPSGSADETLQLQISQLDYDNYTGRLGIGRILNGRIKPGQTVAVMN					
g151	DMRPLFDITLKYTPAPSGSADEPLQLQISQLDYDNYTGRLGIGRILNGRIKPGQTVAVMN					
	190	200	210	220	230	240
m151.pep	250	260	270	280	290	300
	HDQQIAQGRINQLLGFKGLERVPLEEAEAGDIVIISGIEDIGIGVTITDKDNPKGLPMLS					
g151	HEQQIAQGRINQLLGFKGLERVPLEEAEAGDIVIISGIEDIGIGVTITDKDNPKGLPMLS					
	250	260	270	280	290	300
m151.pep	310	320	330	340	350	360
	VDEPTLTMDFMVNTSPLAGTEGKFVTSRQIRDRLQKELLTNVALRVEDTADADVFRVSGR					
g151	VDEPTLTMDFMVNTSPLAGTEGKFVTSRQIRDRLQKELLTNVALRVEDTADADVFRVSGR					
	310	320	330	340	350	360
m151.pep	370	380	390	400	410	420
	GELHLTILLENMRREGYELAVGKPRVVYRDIDGQKCEPYENLTVDPDDNQAVMEELGR					
g151	GELHLTILLENMRREGYELAVGKPRVVYRDIDGQKCEPYENLTVDPDDNQAVMEELGR					
	370	380	390	400	410	420
m151.pep	430	440	450	460	470	480
	RRGELTNMESDGNRTRLEYHIPARGLIGFQGEFMTLTRGVGLMSHVDDYAPVKPDMPG					
g151	RRGELTNMESDGNRTRLEYHIPARGLIGFQGEFMTLTRGVGLMSHVDDYAPVKPDMPG					
	430	440	450	460	470	480
m151.pep	490	500	510	520	530	540
	RHNGVLVSQEQGEAVAYALWNLEDGRMFVSPNDKIYEGMIIGIHSRDNDLVNPLKGKK					
g151	RHNGVLVSQEQGEAVAYALWNLEDGRMFVSPNDKIYEGMIIGIHSRDNDLVNPLKGKK					
	490	500	510	520	530	540
m151.pep	550	560	570	580	590	600
	LTNIRASGTDEAVRLTTPIKLTLEGAVEFIDDELVEITPQSIRLRKRYLSELERRRHFK					
g151	LTNIRASGTDEAVRLTTPIKLTLEGAVEFIDDELVEITPQSIRLRMYLSELERRRHFK					
	550	560	570	580	590	600

m151.pep KLDX  
 ||||  
 g151 KLDX

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 617>:

a151.seq  
 1 ATGAAACAAA TCCGCAACAT CGCCATCATC GCCACGTCG ACCACGGCAA  
 51 AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA  
 101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA  
 151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA  
 201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GCGGCGAAG  
 251 TAGAGCGAGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG  
 301 CAGGAAGGCC CGATGCCGCA AACCCTTTTC GTGACCAAAA AAGCCTTGGC  
 351 TTTGGGGCTG AAACCGATTG TCGTCATCAA TAAAATCGAC AAACCGTCCG  
 401 CCCGTCCGAG CTGGGTCATC GACCAAACCT TCGAGCTGTT CGACAACCTG  
 451 GGCGCGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTATG CTTCGGGTCT  
 501 GTCCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC  
 551 CGCTGTTTCA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG  
 601 GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC  
 651 CGGCCGCTC GGTATCGGTC GTATCTTGAA CGGACGTATC AAGCCCGGTC  
 701 AAGTTGTTGC CGTCATGAAC CACGATCAAC AAATCGCCCA AGGCCGCATC  
 751 AACCAGCTTT TGGGTTTCAA AGGTTTAGAA CGCGTGCCGC TTGAAGAAGC  
 801 CGAAGCCGGC GACATCGTGA TTATTTCCGG TATTGAAGAC ATCGGCATCG  
 851 GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCCTGCC GATGTTGAGC  
 901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGTT  
 951 GGCAGGTACG GAAGGCAAAT TCGTAACCAG CCGCCAAATC CGGCACCGCC  
 1001 TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCA  
 1051 GATGCCGACG TGTTCGCGT ATCCGGGCGC GCGAGCTGC ACCTGACCAT  
 1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC  
 1151 CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA  
 1201 AACCTGACCG TGGACGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA  
 1251 ACTCGGCCCG CGCCGTGGCG AACTGACTAA TATGGAAGC GACGGCAACG  
 1301 GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGCTTC  
 1351 CAAGGCGAAT TTATGACCCT GACGCGCGGG GTCGGGCTGA TGAGCCACGT  
 1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCTGGC CGCCACAACG  
 1451 GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTTG  
 1501 AATCTGGAAG ACCGCGGCCG TATGTTCTGA TCGCCCAACG ACAAATCTA  
 1551 CGAAGGTATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGCTCA  
 1601 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC  
 1651 GAAGCCGTTT GCCTGACCAC GCCGATTAAG CTGACGCTGG AAGGTGCGGT  
 1701 CGAGTTTATC GACGATGATG AGCTGGTAGA AATCACGCCG CAATCCATCC  
 1751 CTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCCG CCATTTCAAA  
 1801 AAGCTAGATT GA

This corresponds to the amino acid sequence <SEQ ID 618; ORF 151.a>:

a151.pep  
 1 MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE  
 51 RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVLLVDA  
 101 QEGPMPQTRF VTKKALALGL KPIVINKID KPSARPSWVI DQTFELFDNL  
 151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDLIL KYTPAPSGSA  
 201 DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQVVAVMN HDQQIAQGRI  
 251 NQLLGFKGLE RVPLEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS  
 301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLLQKELL NVALRVEDTA  
 351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGKCPEPYE  
 401 NLTVDVPDDN QGAVMEELGR RRGELTNMES DNGNRTRLEY HIPARGLIGF  
 451 QGEFMTLTRG VGLMSHVFD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW  
 501 NLEDRGRMFV SPNDKIYEGM IIGIHSRND LVVNPLKGKK LTNIRASGTD  
 551 EAVRLTPIK LTLEGAVEFI DDDELVEITP QSIRLRKRYL SELERRRHFK  
 601 KLD\*

m151/a151 99.8% identity in 603 aa overlap

10 20 30 40 50 60

m151.pep	MKQIRNIAIIAHVDHGKTTLV	DQLLRQSGTFRANQQVDERV	MDSNDLEKERGITILAKNT		
a151	MKQIRNIAIIAHVDHGKTTLV	DQLLRQSGTFRANQQVDERV	MDSNDLEKERGITILAKNT		
	10	20	30	40	50
	60				
m151.pep	AIDYEGYHINIVDTPGHADFG	GEVERVLGMVDCVLLVDAQEG	PMPQTRFVTKKALALGL		
a151	AIDYEGYHINIVDTPGHADFG	GEVERVLGMVDCVLLVDAQEG	PMPQTRFVTKKALALGL		
	70	80	90	100	110
	120				
m151.pep	KPIVVINKIDKPSARPSWVIDQ	TFFELFDNLGATDEQLDFPIV	YASGLSGFAKLEETDES		
a151	KPIVVINKIDKPSARPSWVIDQ	TFFELFDNLGATDEQLDFPIV	YASGLSGFAKLEETDES		
	130	140	150	160	170
	180				
m151.pep	DMRPLFDITILKYTPAPSGSA	DETLLQLQISQLDYDNYTGR	LIGRILNGRIKPGQTVAVMN		
a151	DMRPLFDITILKYTPAPSGSA	DETLLQLQISQLDYDNYTGR	LIGRILNGRIKPGQTVAVMN		
	190	200	210	220	230
	240				
m151.pep	HDQQIAQGRINQLLGFKGLER	VPLEEAEAGDIVIISGIEDIG	IGVTITDKDNPKGLPMLS		
a151	HDQQIAQGRINQLLGFKGLER	VPLEEAEAGDIVIISGIEDIG	IGVTITDKDNPKGLPMLS		
	250	260	270	280	290
	300				
m151.pep	VDEPTLTMDFMVNTSPLAGTE	GKFVTSRQIRDLQKELLTNV	ALRVEDTADADVFRVSGR		
a151	VDEPTLTMDFMVNTSPLAGTE	GKFVTSRQIRDLQKELLTNV	ALRVEDTADADVFRVSGR		
	310	320	330	340	350
	360				
m151.pep	GELHILTILLENMRREGYELAV	GKPRVVYRDIDGQKCEPYEN	LTVDVPDDNQAVMEELGR		
a151	GELHILTILLENMRREGYELAV	GKPRVVYRDIDGQKCEPYEN	LTVDVPDDNQAVMEELGR		
	370	380	390	400	410
	420				
m151.pep	RRGELTNMESDGNGRTRLEYH	I PARGLIGFQGEFMTLTRGV	GLMSHVDDYAPVKPDMPG		
a151	RRGELTNMESDGNGRTRLEYH	I PARGLIGFQGEFMTLTRGV	GLMSHVDDYAPVKPDMPG		
	430	440	450	460	470
	480				
m151.pep	RHNGVLVSQEQGEAVAYALWN	LEDGRGMFVSPNDKIYEGMI	IIGIHSRDNDLVVNPLKGKK		
a151	RHNGVLVSQEQGEAVAYALWN	LEDGRGMFVSPNDKIYEGMI	IIGIHSRDNDLVVNPLKGKK		
	490	500	510	520	530
	540				
m151.pep	LTNIRASGTDEAVRLTTPIKL	TLEGAVEFIDDELVEITPQSI	RLRKRYLSELEERRRHFK		
a151	LTNIRASGTDEAVRLTTPIKL	TLEGAVEFIDDELVEITPQSI	RLRKRYLSELEERRRHFK		
	550	560	570	580	590
	600				
m151.pep	KLDX				
a151	KLDX				

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 619>:

g152.seq

```

1  ATGAAAAaca aaACCaaagt ctgGGacttc cCaccgcgc tttTCCactG
51  GctgcttgCC gCATCCctgc CCTTTATGTG gtatagCGCA AAAGCCGCGC
101 GcgataTGCT GcaatgGCAC ACGCGCTCGT GGCTGCTCGT CCTTTTCCTG

```



```

151 CTCGTATTCC GCCTCTGCTG GGGCATTGTT GGCAGcgATA CCGCCCGTTT
201 CTCccgTtTC GTCCGAGGTT GGGCAGGTAT ACGCGGCTAT CTGAAAAAcg
251 gCATTCCCCG ACAtatcCAG CCGGACACA ACCCTTGCG CGCACTgatg
301 gtcGTTGCGC TTTTGgcccgc cgtcTCATTT CAagtcggcA CGGGGCTTTT
351 Tgccgccaat gaaaacacct tcagcaCCaA cggctacctc aaccatttgg
401 tttccgaaca tacgGGCAGC CTTATACGGA AAATCCACCT CAACTTTTTC
451 AAGCTGCTCG CCGTTTTTTC CGCAGTCCAC ATCGCCGCCG TCGCCGCATA
501 CCGCATATTC AAAAAGAAAA ACCTCGTCCG CCCGATGATA ACCGGCTTCA
551 AATACATCGA AGGCAAAACC TCAATCCGCT TTGCCGGCAA AGCCGCGCTT
601 GCCGCCGCAT TATCGGTTGC CGCGCTTGCC GCAGCCGCCA TCCTGCTCCT
651 GTCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 620; ORF 152.ng>:

g152.pep

```

1 MKNKTKVWDF PTRLFWLLA ASLPFMWYSA KAGGDMLOWH TRVGLLVFLFL
51 LVFRLCWGIW GSDTARFSRF VRGWAGIRGY LKNGIPEHIQ PGHNPLGALM
101 VVALLAAVSF QVGTGLFAAN ENTFTNGYL NHLVSEHTGS LIRKIHLNFF
151 KLLAVFSAVH IAAVAAYRIF KKKNLVRPMI TGFKYIEGKT SIRFAGKAAL
201 AAALSVAALA AAAILLLS*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 621>:

m152.seq

```

1 ATGAAAAACA AAACCAAAGT CTGGGACCTC CCCACCCGCC TTTTCCACTG
51 GCTGCTTGCC GCGTCCCTGC CCTTTATGTG GTATAGCGCG AAAGCCGGCG
101 GCGATATGCT GCAATGGCAC ACGCGCGTCG GGCTGTTCTG CTTTTCCTG
151 CTCGTATTTC GCCTCTGCTG GGGCATTGTT GGCAGCGATA CCGCCCGTTT
201 TTCCCGTTTC GTCCAAGGCT GGGCAGGCAT ACGCGGCTAT CTGAAAAACG
251 GTATTCCCGA ACACATCCAG CCGGACACA ACCCTTGCG CGCACTGATG
301 GTCGTTGCGC TTTTGCGCCG CGTGTCCTTC CAAGTCGGCA CCGGGCTTTT
351 TGCCGCCGAT GAAAACACCT TCAGCACCAA CGGCTACCTC AACCATTGTTG
401 TTTCCGAACA TACGGGCAGC CTTATGCGGA AAATCCACCT CAACTTTTTC
451 AAGCTGCTCG CCGTTTTTTC TGCAATCCAC ATCGCCGCCG TCGCCGCATA
501 CCGCGTATTC AAAAAGAAAA ACCTCATCCT CCCGATGATA ACCGGCTTCA
551 AATACATCGA AGGCAAAACC TCAATCCGCT TTGCAGGCAA AGCCGCGCTT
601 GCCGCCGCAT TATCGGTTGC CTCGCTTGCC GCAGCCGCCA TCCTGCTCCT
651 GTCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 622; ORF 152>:

m152.pep

```

1 MKNKTKVWDL PTRLFWLLA ASLPFMWYSA KAGGDMLOWH TRVGLFVLFL
51 LVFRLCWGIW GSDTARFSRF VQGWAGIRGY LKNGIPEHIQ PGHNPLGALM
101 VVALLAAVSF QVGTGLFAAD ENTFTNGYL NHLVSEHTGS LMRKIHLNFF
151 KLLAVFSAIH IAAVAAYRVF KKKNLILPMI TGFKYIEGKT SIRFAGKAAL
201 AAALSVAALA AAAILLLS*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 152 shows 95.4% identity over a 218 aa overlap with a predicted ORF (ORF 152.ng) from *N. gonorrhoeae*:

m152/g152

```

          10      20      30      40      50      60
m152.pep  MKNKTKVWDLPTTRLFWLLAASLPFMWYSAKAGGDMLOWHTRVGLFVLFLFLLVFRLCWGIW
          |||||:|||||
g152      MKNKTKVWDFPTRLFWLLAASLPFMWYSAKAGGDMLOWHTRVGLLVFLFLLVFRLCWGIW
          10      20      30      40      50      60

          70      80      90     100     110     120
m152.pep  GSDTARFSRFVQGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAD
          |||||:|||||
g152      GSDTARFSRFVRGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAN
          70      80      90     100     110     120

```

	130	140	150	160	170	180
m152.pep	ENTFSTNGYLNHLVSEHTGSLMRKIHLNFFKLLAVFSAIHIAAVAAYRVFKKKNLILPMI					
g152	ENTFSTNGYLNHLVSEHTGSLIRKIHLNFFKLLAVFSAVHIAAVAAYRIFKKNLVRPMI					
	130	140	150	160	170	180
	190	200	210	219		
m152.pep	TGFKYIEGKTSIRFAGKAALAAALSVAASLAAAAILLLSX					
g152	TGFKYIEGKTSIRFAGKAALAAALSVAALAAAAILLLSX					
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 623>:

```

a152.seq
1  ATGAAAAACA AAACCAAAGT CTGGGACTTC CCCACCCGCC TTTTCCACTG
51  GCTGCTTGCC GCATCCCTAC CCTTTATGTG GTATAGCGCG AAAACCGGCG
101 GCGATATGCT GCAATGGCAC ACGCGCGTCG GGCTGTTTAT CCTTTTCCTG
151 CTCGTATTCC GCCTCTGCTG GGGCATTGCG GGCAGCGATA CCGCCCGTTT
201 CTCCCGTTTC GTCCGCGGAT GGTCCGGTAT CAGAGAGTAT ATGAAAAACG
251 GTATTCCCGA ACACGTCCAA CCCGGACACA ACCCCTTGGG CGCACTGATG
301 GTCGTTGCGC TTTTGGCCGC CGTGTCGTTC CAAGTCGGCA CAGGGCTTTT
351 TGCCGCCGAT GTAAACACCT TCAGCACCAA CGGCTACCTC AACCATTGCG
401 TTTCCGAACA TACGGGCAGC CTTATGCGGA AAATCCATCT CAACTTTTTC
451 AAACGTCTCG CCGTTTTTTC CGCAGTCCAC ATCGCCGNCG TCGCCGCATA
501 CCGCGTGTTC AAAAAGAAAA ACCTCGTCCT CCCGATGATA ACCGGCTTCA
551 AATACATCGA AGGCAAAACC TCAATCCGCT TTGCCGGCAA AGCCGCGCTT
601 GCCGCCGCAT TATCGGTTGC CGCGCTTGCC GCAGCCGCCA TCCTGCTCCT
651 GTCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 624; ORF 152.a>:

```

a152.pep
1  MKNKTKVWDF PTRLFHWLLA ASLPFMWYSA KTGDMQLQWH TRVGLFILFL
51  LVFRLCWGIW GSDTARFSRF VRGWSGIREY MKNGIPEHVQ PGHNPLGALM
101 VALLAAVSF QVGTGLFAAD VNTFSTNGYL NHLVSEHTGS LMRKIHLNFF
151 KLLAVFSAVH IAXVAAYRVF KKKNLVLPMI TGFKYIEGKT SIRFAGKAAL
201 AAALSVAALA AAAILLLS*

```

m152/a152 94.0% identity in 218 aa overlap

	10	20	30	40	50	60
m152.pep	MKNKTKVWDLPTLRFHWLLAASLPFMWYSAKAGGDMQLQWHTRVGLFVLFLVLRFCWGIW					
a152	MKNKTKVWDFPTLRFHWLLAASLPFMWYSAKTGGDMQLQWHTRVGLFILFLVLRFCWGIW					
	10	20	30	40	50	60
	70	80	90	100	110	120
m152.pep	GSDTARFSRFVQGWAGIRGYLKNIGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAD					
a152	GSDTARFSRFVRGWSGIREYMKNGIPEHVQPGHNPLGALMVVALLAAVSFQVGTGLFAAD					
	70	80	90	100	110	120
	130	140	150	160	170	180
m152.pep	ENTFSTNGYLNHLVSEHTGSLMRKIHLNFFKLLAVFSAIHIAAVAAYRVFKKKNLILPMI					
a152	VNTFSTNGYLNHLVSEHTGSLMRKIHLNFFKLLAVFSAVHIAAXVAAYRVFKKKNLVLPMI					
	130	140	150	160	170	180
	190	200	210	219		
m152.pep	TGFKYIEGKTSIRFAGKAALAAALSVAASLAAAAILLLSX					
a152	TGFKYIEGKTSIRFAGKAALAAALSVAALAAAAILLLSX					
	190	200	210			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 625>:

g153.seq  
 1 atgggggtttg cttacAgtat gacgtatatc gaggtCGGGA taccggaggc  
 51 ggcattccgtc ctttCgctGC CCGAGATgat gcgcctgatG GTGTTtCagg  
 101 attATGGTTT TttggcCGAA GTGATGTTTG TGctgaCTTT cGGCGcgcCG  
 151 GTTCTGTTtC TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA  
 201 ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA  
 251 GGCAGGCGAT GATGGTGGAT GTGTTTTTTG TTCCACTCT GGTGGCGTAT  
 301 ATCAAGCTCT CGTCTGTGGC AAAGGTTTCG TCGGGGCCG CGTTTTATCT  
 351 GATGTTCCGC CTGTCGGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC  
 401 AGCATTGGGT GTATTTCCAA ATCGGGCGGC TGACGGGGAA TAATGCGGTT  
 451 CAGACGGCAT CGGAAGGCAA AACCTGTGTC AGCCGCTGCC TGTATTTccg  
 501 cgacAGTgcc gaatccCCCT GCGGGGTGTg cgCGGcgggA CTgtacggcg  
 551 gacggccgaa aagtCTGAGt atttCgtCGG CGTTTCTgac ggcggcggtT  
 601 GTTTTGTATT TCCctgCcaa TATCctgccc attaTGattt cgtccAATCc  
 651 tgccgccacg GAGGcCAACA CCATCTTTAG CGGCATCGCT TATATGTGGG  
 701 ACGagggcgA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG  
 751 GTGCCGGTGC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGGCGGCACG  
 801 VLYFPANILP IMISSNPAAT EANTIFSGIA YMWDEGDRLI AAVIFSASIL  
 851 CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTGTG GATTATTATT  
 901 TTGATGTGTT CGTTCCacac TTATGCCGCG CGCGTCATTc CGGGCAGTGC  
 951 GGCAGTCTAT TTCTGCCTGG TCGTGATTTT GACGATGCTG TCCGCCTATT  
 1001 ATTTTCGACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT  
 1051 TTCAACGAAA CGGAAAAATA TGACTGA

This corresponds to the amino acid sequence <SEQ ID 626; ORF 153.ng>:

g153.pep  
 1 MGFAYSMTYI EVGIPEAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP  
 51 VLFLLLCLYV YAALIRKQAY PALRLATRV M VRLRQAMVD VFFVSTLVAY  
 101 IKLSSVAKVR FGPAFYLMFA LSVMLIRTSV SVPQHWVYFQ IGRLTGNNAV  
 151 QTASEGKTCC SRCLYFRDSA ESPCGVCGAE LYGGPRPKSLs ISSAFLTAAY  
 201 VLYFPANILP IMISSNPAAT EANTIFSGIA YMWDEGDRLI AAVIFSASIL  
 251 VPVLKIAAMS VLIAAARFAL PAGAKKLSHL YRITEAVGRW SMIDIFVIII  
 301 LMCSFHTYAA RVIPGSAAVY FCLVVILTML SAYYFDPRL L WDKRASDGIA  
 351 FNETEKYD\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 627>:

m153.seq  
 1 ATGGCGTTTG CTTACGGTAT GACGTATATC GAGGTCGGGA TACCGGGTGC  
 51 GGCATCCGTC CTTTCGCTGC CCGAGATGAT GCGCCTGATG GTGTTTCAGG  
 101 ATTATGGTTT TTTGGCCGAA GTGATGTTTG TGCTGACTTT CGGCGCGCCG  
 151 GTTCTGTTTC TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA  
 201 ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA  
 251 GACAGGCGAT GATGGTGGAT GTGTTTTTTG TTCCACTTT GGTGGCGTAT  
 301 ATCAAGCTCT CGTCTGTGGC AGAGGTTTCG TCGGGGCCG CGTTTTATCT  
 351 GATGTTCCGC CTGTCAGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC  
 401 AGCATTGGGT GTATTTTCAA ATCGGGCGGC TGACGGGGGA TAATGCGGTT  
 451 CAGACGGCAT CGGAAGGTAA AACCTGTGTC AGCCGCTGCC TGTATTTCCG  
 501 CGACAGTGCC GAATCCCCCT GCGGCGTGTG CGGTGCGGAA CTGTACCGCC  
 551 GACGGCCGAA AAGTCTGAGT ATTTCTGCGG CGTTTCTGAC GGCGGCGGTT  
 601 ATTTTGTATT TCCCTGCCAA TATCCTGCCG ATTATGATTT CGTCCAATCC  
 651 TGCCGCCACG GAGGTCAATA CCATCCTTAA CGGCATCGCT TATATGTGGG  
 701 ACGAGGGCGA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG  
 751 GTGCCGGTAC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGTCCGCCCG  
 801 CTTCTGCTTG CCAACGGGTG CAAAGAAATT GTCGCACCTC TACCGCATCA  
 851 CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTGTG GATTATTATT  
 901 TTGATGTGTT CGTTCCACAC TTATGCCGCG CGCGTCATTc CGGGCAGTGC  
 951 GGCAGTCTAT TTCTGCCTGG TCGTGATTTT GACGATGCTG TCCGCCTATT  
 1001 ATTTTCGACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT  
 1051 TTCAATGAAA CGGAAAAACA TGACTGA

This corresponds to the amino acid sequence <SEQ ID 628; ORF 153>:

m153.pep  
 1 MAFAYGMTYI EVGIPGAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP  
 51 VLFLLLCLYV YAALIRKQAY PALRLATRV M VRLRQAMVD VFFVSTLVAY

101 IKLSSVAEVR FGPAFYLMFA LSVMLIRTSV SVPQHWVYFQ IGRLTGDNAV  
 151 QTASEGKTCC SRCLYFRDSA ESPCGVCGAE LYRRRPKSL ISSAFLTA  
 201 ILYFPANILP IMISSNPAAT EVNTILNGIA YMWDEGDRLI AAVIFSASIL  
 251 VPVLKIAAMS VLIASARFAL PTGAKKLSHL YRITEAVGRW SMIDIFVIII  
 301 LMCSFHYYAA RVIPGSAAVY FCLVVILTML SAYYFDPRL WDKRASDGIA  
 351 FNETEKHD\*

m153 / g153 96.1% identity in 358 aa overlap

m153.pep	MAFAYGMTYIEVGIPGAASVLSLPEMMRLMVFDYGFLEVMFVLTFGAPVLFLLCLYV
g153	MGFAYSMTYIEVGIPGAASVLSLPEMMRLMVFDYGFLEVMFVLTFGAPVLFLLCLYV
m153.pep	YAAIRKQAYPALRLATRVMLRQAMMVDVFFVSTLVAYIKLSSVAEVRFGPAFYLMFA
g153	YAAIRKQAYPALRLATRVMLRQAMMVDVFFVSTLVAYIKLSSVAKVRFGPAFYLMFA
m153.pep	LSVMLIRTSVSVPOHWVYFQIGRLTGDNAVQTASEGKTCCSRCLYFRDSAESPCGVCGAE
g153	LSVMLIRTSVSVPOHWVYFQIGRLTGNNAVQTASEGKTCCSRCLYFRDSAESPCGVCGAE
m153.pep	LYRRRPKSLSISSAFLTAAVILYFPANILPIMISSNPAATEVNTILNGIAYMWDEGDRLI
g153	LYGGRPKSLSISSAFLTAAVILYFPANILPIMISSNPAATEANTIFSGIAYMWDEGDRLI
m153.pep	AAVIFSASILVPVLKIAAMSVLIASARFALPTGAKKLSHLRYRITEAVGRWSMIDIFVIII
g153	AAVIFSASILVPVLKIAAMSVLIAAARFALPAGAKKLSHLRYRITEAVGRWSMIDIFVIII
m153.pep	LMCSFHYYAARVIPGSAAVYFCLVVILTMLSAYYFDPRLWDKRASDGIAFNETEKHDX
g153	LMCSFHYYAARVIPGSAAVYFCLVVILTMLSAYYFDPRLWDKRASDGIAFNETEKYDX

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 629>:

a153.seq

1 ATGGCGTTTG CTTACGGTAT GACGTATATC GAGGTCGGGA TACCGGGTGC  
 51 GGCATCCGTC CTTTCGCTGC CCGAGATGAT GCGCCTGATG GTGTTTCAGG  
 101 ATTATGGTTT TTTGGCCGAA GTGATGTTTG TGCTGACCTT CGGCGCGCCG  
 151 GTTCTGTTTC TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA  
 201 ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA  
 251 GACAGGCGAT GATGGTGGAT GTGTTTTTTG TTTCCACTTT GGTGGCGTAT  
 301 ATCAAGCTCT CGTCTGTGGC AGAGGTTTCG TTCGGATCGG CGTTTTATCT  
 351 GATGTTTCGG CTGTCGGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC  
 401 AGCATTGGGT GTATTTTCAA ATCGGCGCGC TGACGGGGGA TAATGCGGTT  
 451 CAGACGGCAT CGGAAGGTAA AACCTGTTGC AGCCGCTGCC TGTATTTCCG  
 501 CGACAGTGCC GAATCCCCCT GCGGCGTGTG CGGTGCGGAA CTGTACCGCC  
 551 GACGGCCGAA AAGTCTGAGT ATTCGTCGG CGTTTCTGAC GGCGCGGTT  
 601 ATTTTGTATT TCCCTGCCAA TATCCTGCCG ATTATGATT CGTCCAATCC  
 651 TGCCGCCACG GAGGTCAATA CCATCCTTAA CGGCATCGCT TATATGTGGG  
 701 ACGAGGGCGA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG  
 751 GTGCCGGTAC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGTCCGCCCG  
 801 CTTTCGCTTTG CCAACGGGTG CAAAGAAATT GTCGCACCTC TACCGCATCA

```

851 CCGAAGCGGT CCGCCGCTGG TCGATGATTG ATATTTTGT GATTATTATT
901 TTGATGTGTT CGTTCCACAC TTATGCCGCG CGCGTCATTG CCGGCAGTGC
951 GGCAGTCTAT TTCTGCCTGG TCGTGATTCT GACGATGCTG TCCGCCTATT
1001 ATTTTCACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT
1051 TTCAATGAAA CGGAAAAACA TGACTGA

```

This corresponds to the amino acid sequence <SEQ ID 630; ORF 153.a>:

```

a153.pep
  1 MAFAYGMTYI EVGIPGAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
 51 VLFLLCLYV YAALIRKQAY PALRLATRV MRLQAMMVD VFFVSTLVAY
101 IKLSSVAEVR FGSAYFLMFA LSVMLIRTSV SVFQHWVYFQ IGRLTGDNV
151 QTASEGKTCC SRCLYFRDSA ESPCGVCGAE LYRRRPKSL SISAFLTA
201 ILYFPANILP IMISSNPAAT EVNTILNGIA YMWDEGDRLI AAVIFSASIL
251 VPVLKIAAMS VLIASARFAL PTGAKKLSHL YRITEAVGRW SMIDIFVIII
301 LMCSFHTYAA RVIPGSAAVY FCLVVILTML SAYYFDPRL WDKRASDGIA
351 FNETEKHD*

```

m153/a153 99.7% identity in 358 aa overlap

	10	20	30	40	50	60
m153.pep	MAFAYGMTYIEVGIPGAASVLSLPEMMRLMVFQDYGFLAEVMFVLTFGAPVLFLLCLYV					
a153	MAFAYGMTYIEVGIPGAASVLSLPEMMRLMVFQDYGFLAEVMFVLTFGAPVLFLLCLYV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m153.pep	YAALIRKQAYPALRLATRVMLRQAMMVDVFFVSTLVAYIKLSSVAEVRFGPAFYLMFA					
a153	YAALIRKQAYPALRLATRVMLRQAMMVDVFFVSTLVAYIKLSSVAEVRFGSAFYLMFA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m153.pep	LSVMLIRTSVSVFQHWVYFQIGRLTGDNVQTASEGKTCCSRCLYFRDSAESPCGVC					
a153	LSVMLIRTSVSVFQHWVYFQIGRLTGDNVQTASEGKTCCSRCLYFRDSAESPCGVC					
	130	140	150	160	170	180
	190	200	210	220	230	240
m153.pep	LYRRRPKSLSSISAFLTAIVILYFPANILPIMISSNPAATEVNTILNGIAYMWDEGDR					
a153	LYRRRPKSLSSISAFLTAIVILYFPANILPIMISSNPAATEVNTILNGIAYMWDEGDR					
	190	200	210	220	230	240
	250	260	270	280	290	300
m153.pep	AAVIFSASILVPVLKIAAMSVLIASARFALPTGAKKLSHLRYRITEAVGRWSMIDIFV					
a153	AAVIFSASILVPVLKIAAMSVLIASARFALPTGAKKLSHLRYRITEAVGRWSMIDIFV					
	250	260	270	280	290	300
	310	320	330	340	350	359
m153.pep	LMCSFHTYAAARVIPGSAAVYFCLVVILTMLSAYYFDPRLWDKRASDGIAFNETEKH					
a153	LMCSFHTYAAARVIPGSAAVYFCLVVILTMLSAYYFDPRLWDKRASDGIAFNETEKH					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 631>:

```

g154.seq
  1 ATGACTGACA ACAGCCCTCC TCCAAACGGA CACGCTCAAG CACGCGTCCG
 51 CAAAAACAAC accttctctCT CCGCCGCTCTG GCTGGTCCCG CTGATCGCGC
101 TGATTGCCGG CGGCTGGCTT TGGGTAAAGG AAATCCGCAA CAGGGGGCCT
151 GTGGTTACGC TCTTGATGGA CAGCGCGGAA GGCATCGAAG TCAACAATAC
201 GGTCATTAAG GTATTGAGCA TCGATGTCGG ACGCGTTACC CGAATCAAAC
251 TGCGCGACGA CAAAAAGGC GTGGAAGTTA CTGCCCACT CAATGCGGAC
301 GTATCCGCGC TCATCCGCGC CGATACCCAG TTTTGGGTGG TCAAGCCGCG
351 TATCGACCAA AGCGGcgtaa CCGGTTTGGG TACGCTGCTT TCGGGTTCGT

```

```
401 ACATCGCTTT TACACCCGGC AAAAGCGGCG AGGCAAAAGA CGTGTTCCTAA
451 GTGCAGGACA TTCCGCCCGT TACCGCCATC GGGCAAAAGcG GGCTGCGCTT
501 GAATTGATT GGTAAAAACG AccgCATCCT CAACGTcaaC AGCCCTGTTT
551 TGTATGAAAA CTTTATGGTC GGGCAAATCG AAAGCGCGCA TTTCGAcCCG
601 TCCGACCAAA GCGTGCATTA CACCATCTTC ATCCAAAGCC CCAACGACAA
651 ACTGATTCAT TCCGCCAGCC GTTTTGGCT GGAAAGCGGC ATCAATATCG
701 AAACCACAGG CAGCGGCATC AAACCTCAATT CCGCCCTCT GCCTGCCCTG
751 CTGTCAGGCG CGATTTCATT TGATTCGCCG AAAACCAAAA ACAGTAAAAA
801 CGTCAAAAGC GAGGACAGCT TCACGCTTTA CGACAGCCGC AGCGAAATCG
851 CCAACCTGCC TGACGACCGC TCGCTGTACT ACACCGCGTT TTTCAAACAA
901 TCCGTGCGCG GACTGACCGT cggTTCGCCT GTcgaATACA AAGGGCtgaA
951 TGTCggCATG GTTTCGGATG TCCCTTATTT TGACCGCAAt gacagCCTGC
1001 ACCTgtTTGA aaacggctgg aTTcccGtac gCATCCGCAT cgagccTTCC
1051 CGTTTGGAAA TCAATGCCGA CGAGCAAAGC AAAGAGCATT GGAACAACA
1101 ATTCCAGACG GCCTTAAACA AAGGCTGAC CGCCACCATC TCCAGCAACA
1151 ACCTGCTGAC CGGCGGCAAA ATGATTGAGT TGAACGATCA GCCTTCCGCC
1201 TCGCCCAAGC TCGGACCGCA TACCGTTTAT GCAGGCGATA CCGTCATCGC
1251 CACACGGGCG GCGGTTTGG ATGACTTGCA GGTCAAATTG GCGGATTGTC
1301 TGGACaaatT CAACAATCTG CCATTggata aaACCGTGC CGAATTGAAC
1351 GGCTCGCTCG CCGAACTCAA GTCCGCACTC AAATCCGCCA ATGCCGCCCT
1401 AAGCTCCATT GacaAACTGG TCGgcaaTCC GCAGACGCAA AACATCCCGA
1451 ACGAACTGAA CCAAACTCTG AAAGAGTTGC GCATAACCTT GCAAGGCGTA
1501 TCGcctCAAT CGCCTATCTa cgGagacgta caAAATAcgc tgCaAAGTTT
1551 GGACAAAACC TTAaAgacg TtcaACCGT CATTAACT TTAaAGAAa
1601 aacCCaaCgc actGATTtTc aacaACAGCA GCAAAGAccc tATCCCGAAA
1651 GGAAGCCGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 632; ORF 154.ng>:

g154.pep

```
1 MTDNSPPPNG HAQARVRKNN TFLSAVWLV LIALIAGGWL WVKEIRNRGP
51 VVTLMDSAE GIEVNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD
101 VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTL SSGYIAFTPG KSGEAKDVFO
151 VQDIPVTAI QSGRLRLNLI GKNDRIILNVN SPVLYENFMV GQIESAHFDP
201 SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETGSGI KLNSAPLPAL
251 LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEIANLPDDR SLYYTAFFKQ
301 SVRGLTVGSP VEYKGLNVGM VSDVPYFDRN DSLHLFENGW IPVRIRIEPS
351 RLEINADEQS KEHWKQQFQT ALNKGLTATI SSNNLLTGSK MIELNDQPSA
401 SPKLRPHTVY AGDTVIAIRG GGLDDLQVKL ADLLDKFNFL PLDKTVAELN
451 GSLAELKSAL KSANAALSSI DKLVGNPQTQ NIPNELNQT LKELRITLQGV
501 SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKFNALIF NNSKDPPIK
551 GSR*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 633>:

m154.seq

```
1 ATGACTGACA ACAGCCCTCC TCCAAACGGA CACGCCAAG CACGCGTCCG
51 CAAAACAAAC ACCTTCCTCT CTGCCGTCTG GCTGGTCCG CTGATCGCGC
101 TGATTGCCCG CGGCTGGCTT TGGGTAAAG AAATCCGCA CAGGGGGCCT
151 GTGGTTACGC TCTTGATGGA CAGCGCGGAA GGCATTGAG TCAACAATAC
201 GGTCATCAAA GTATTGAGCA TCGATGTCGG ACGCGTTACC CGAATCAAAC
251 TCGCGACGCA CCAAAAAGGC GTGGAAGTAA CCGCCCACT CAATGCGGAC
301 GTATCCGGCC TCATCCGCAG CGATACCCAG TTTTGGGTGG TCAAGCCGCG
351 TATCGACCAA AGCGGCGTAA CCGGTTTGGG TACGCTGCTT TCGGGTTCGT
401 ACATCGCCTT TACACCCGGC AAAAGCGACG AGGCAAAAGA CGTGTTCCTAA
451 GTGCAGGACA TTCCGCCCGT TACCGCCATC GGGCAAAGCG GGCTGCGCTT
501 GAATTGATT GGTAAAAACG ACCGCATCCT CAACGTCAAC AGCCCTGTTT
551 TGTATGAAAA TTTTATGGTC GGGCAAGTCG AAAGCGCGCA TTTCGACCCG
601 TCCGACCAAA GCGTGCATTA CACCATCTTC ATCCAAAGCC CCAACGACAA
651 ACTGATTCAT TCCGCCAGCC GTTTTGGCT GGAAAGCGGC ATCAATATCG
701 AAACCACAGG CAGCGGCATC AAACCTCAATT CCGCCCTCT GCCTGCCCTG
751 CTGTCGGGCG CGATTTCATT TGATTCGCCG AAAACCAAAA ACAGTAAAAA
801 CGTCAAAAGC GAAGACAGCT TCACGCTTTA CGACAGCCGC AGCGAAGTCG
851 CCAACCTGCC TGACGACCGC TCGCTGTACT ACACCGCGTT TTTCAAACAA
901 TCCGTGCGCG GCCTGACCGT CGGTTCCGCC GTCAGTACA AAGGGCTGAA
951 TGTCGCGCTG GTTTCGGACG TTCCTTATTT CGACCGCAAC GACAGCCTGC
1001 ACCTGTTTGA AAACGCGTGG ATACCCGTAC GCATCCGCAT TGAACCTTCC
1051 CGTTTGGAAA TCAATGCCGA CGAACAAAGC AAAGAACATT GGAACAACA
1101 ATTTCCAGAC GCCTTAAACA AAGGCTGAC CGCCACCATC TCCAGCAACA
1151 ACCTGCTGAC CGGAAGCAAA ATGATTGAGT TGAACGATCA GCCTTCCGCA
```

```

1201 TCACCTAAGC TGCAGCCGCA TACCGTTTAT GCAGGCGATA CCGTTATCGC
1251 GACCCAGGGC GCGGTTTGG ACGATTTGCA GGTCAAATTG GCGGATTTGC
1301 TGGACAAGTT CGACAACTG CCTTTAGATA AGACGGTTGC CGAATTGAAC
1351 GGTTGCGCTG CCGAGCTCAA ATCCACACTC AAATCTGCCA ATGCCGCCCT
1401 AAGCTCCATC GACAACTGG TCGGCAAACC GCAGACACAA AACATTCCGA
1451 ACGAACTGAA CCAAACCTG AAAGAGTTGC GCACAACCCT GCAAGGCGTA
1501 TCGCCGCAAT CGCTATCTA CGGCGACGTA CAAAATACGC TGCAAAGTTT
1551 GGACAAAAC TTAAGAGACG TTCAACCCGT GATTAATACT TTGAAAGAAA
1601 AACCCAACGC GCTGATTTTC AACAGCAGCA GCAAAGACCC TATCCCGAAA
1651 GGAAGCCGAT AA

```

This corresponds to the amino acid sequence <SEQ ID 634 ORF 154.a>:

```

m154.pep
1  MTDNSPPPNG HAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP
51  VVTLLMDSAE GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD
101 VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ
151 VQDIPPVTAI QSGGLRLNLI GKNDRIILNVN SPVLYENFMV GQVESAHFDP
201 SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI KLNSAPLPAL
251 LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ
301 SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS
351 RLEINAEQSK KEHWKQFQT ALNKGLTATI SSNNLTGSK MIELNDQPSA
401 SPKLRPHTVY AGDTVIATQG GGLDDLQVKL ADLLDKFDKL PLDKTVAELN
451 GSLAEKSTL KSANAALSSI DKLVGKPQTQ NIPNELNQLT KELRTTLQGV
501 SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NSSSKDPIPK
551 GSR*

```

m154 / g154 97.8% identity in 553 aa overlap

	10	20	30	40	50	60
m154.pep	MTDNSPPPNGHAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP VVTLLMDSAE					
g154	MTDNSPPPNGHAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP VVTLLMDSAE					
	10	20	30	40	50	60
m154.pep	GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNADV SGLIRSDTQ FWVVKPRIDQ					
g154	GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNADV SGLIRSDTQ FWVVKPRIDQ					
	70	80	90	100	110	120
m154.pep	GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNADV SGLIRSDTQ FWVVKPRIDQ					
g154	GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNADV SGLIRSDTQ FWVVKPRIDQ					
	70	80	90	100	110	120
m154.pep	SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ VQDIPPVTAI GQSGGLRLNL IGKNDRIILNVN					
g154	SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ VQDIPPVTAI GQSGGLRLNL IGKNDRIILNVN					
	130	140	150	160	170	180
m154.pep	SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ VQDIPPVTAI GQSGGLRLNL IGKNDRIILNVN					
g154	SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ VQDIPPVTAI GQSGGLRLNL IGKNDRIILNVN					
	130	140	150	160	170	180
m154.pep	SPVLYENFMV GQVESAHFDP SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI					
g154	SPVLYENFMV GQVESAHFDP SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI					
	190	200	210	220	230	240
m154.pep	SPVLYENFMV GQVESAHFDP SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI					
g154	SPVLYENFMV GQVESAHFDP SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI					
	190	200	210	220	230	240
m154.pep	KLNSAPLPALLSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ					
g154	KLNSAPLPALLSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ					
	250	260	270	280	290	300
m154.pep	KLNSAPLPALLSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ					
g154	KLNSAPLPALLSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ					
	250	260	270	280	290	300
m154.pep	SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS RLEINAEQSK					
g154	SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS RLEINAEQSK					
	310	320	330	340	350	360
m154.pep	SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS RLEINAEQSK					
g154	SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS RLEINAEQSK					
	310	320	330	340	350	360
m154.pep	KEHWKQFQTALNKGLTATISSNNLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG					
g154	KEHWKQFQTALNKGLTATISSNNLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATRG					
	370	380	390	400	410	420
m154.pep	KEHWKQFQTALNKGLTATISSNNLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG					
g154	KEHWKQFQTALNKGLTATISSNNLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATRG					

	370	380	390	400	410	420
	430	440	450	460	470	480
m154 . pep	GGLDDLQVKLADLLDKFDKPLDKTVAEIINGSLAELKSTLKSANAALSSIDKLVGKPKQTQ					
g154	GGLDDLQVKLADLLDKFNNLPLDKTVAEIINGSLAELKSALKSANAALSSIDKLVGKPNPQTQ					
	430	440	450	460	470	480
	490	500	510	520	530	540
m154 . pep	NIPNELNQTLELRITLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF					
g154	NIPNELNQTLELRITLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF					
	490	500	510	520	530	540
	550					
m154 . pep	NSSSKDPIPKGSRX					
g154	NNSSKDPIPKGSRX					
	550					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 635>:

```

a154 . seq
1  ATGACTGACA ACAGCCCTCC TCCAAACGGA CACGCCCAAG CACGCGTCCG
51  CAAAAACAAC ACCTTCCTCT CTGCCGTCTG GCTGGTTCCG CTGATCGCGC
101 TGATTGCCCG CGGCTGGCTT TGGGTTAAGG AAATCCGCAA CAGGGGGCCT
151 GTGGTTACGC TCTTGATGGA CAGCGCGGAA GGCATTGAGG TCAACAATAC
201 GGTCATCAAA GTATTGAGCA TCGATGTCGG ACGCGTTACC CGAATCAAAC
251 TCGCGGACGA CCAAAAAGGC GTGGAAGTAA CCGCCCAACT CAATGCGGAC
301 GTATCCGGCC TCATCCGCAG CGATACCCAG TTTTGGGTGG TCAAGCCGCG
351 TATCGACCAA AGCGGCGTAA CCGGTTTGGG TACGCTGCTT TCGGGTTCGT
401 ACATCGCCTT TACACCCGGC AAAAGCGACG AGGCAAAAGA CGTGTTCCAA
451 GTGCAGGACA TTCCGCCCGT TACCGCCATC GGGCAAAGCG GGCTGCGCTT
501 GAATTTGATT GGTA AAAACG ACCGCATCCT CAACGTCAAC AGCCCTGTTT
551 TGTATGAAAA CTTTATGGTC GGGCAAGTCG AAAGCGCGCA TTTTCGACCCG
601 TCCGACCAAA GCGTGCAATTA CACCATCTTC ATCCAAAGCC CCAACGACAA
651 ACTGATTCAT TCCGCCAGCC GTTCTGGGCT GGAAAGCGGC ATCAATATCG
701 AAACCACAGG CAGCGGCATC AAACCTCAAT CCGCCCTCT GCCTGCCCTG
751 CTGTGCGGCG CGATTTTCAAT TGATTGCGCG AAAACCAAAA ACAGTAAAAA
801 CGTCAAAAGC GAAGACAGCT TCACGCTTTA CGACAGCCGC AGCGAAGTCG
851 CCAACCTGCC TGATGACCGT TCGCTGTAAT ACACCGCGTT TTTCAAACAA
901 TCCGTGCGCG GACTGACCGT CGGTCGCTT GTCGAGTACA AAGGGCTGAA
951 TGTCGCGGTG GTTTCCGATG TTCCTTATTT CGACCGCAAC GACAGCCTGC
1001 ACCTGTTTGA AAACGGCTGG ATTCCCGTAC GCATCCGTAT TGAGCCTTCC
1051 CGTTTGGAAA TCAATGCCGA CGAACAAAGC AAAGAACATT GGAAACAACA
1101 ATTTTCAGAC GCCTTAAACA AAGGCCTGAC CGCCACCATC TCCAGCAACA
1151 ACCTGCTGAC CGGCAGCAAA ATGATTGAGT TGAACGATCA GCCTTCCGCC
1201 TCGCCCAAGC TCGGACCGCA TACCGTTTAT GCAGGCGATA CCGTTATCGC
1251 GACCCAGGGC GCGGTTTGG ACGATTGCA GGTCAAATTG GCGGATTTCG
1301 TGGACAAGTT CGACAACTG CCTTTAGATA AGACGGTTGC CGAATTGAAC
1351 GGTTTCGCTT CCGAGCTCAA ATCCCACTC AAATCTGCCA ATGCCGCCCT
1401 AAGCTCCATC GACAACTGG TCGGCAAACC GCAGACACAA AACATTCCGA
1451 ACGAACTGAA CCAAACCTG AAAGAGTTGC GCACAACCTT GCAAGGCGTA
1501 TCGCCTCAAT CGCCTATCTA CCGCGACGTA CAAATACGCG TGCAAGTTT
1551 GGACAAAACC TTAAAGACG TTCAACCCGT CATTAACACT TTGAAAGAAA
1601 AACCCAACGC GCTGATTTTC AACAGCAGCA GCAAAGACCC TATCCCGAAA
1651 GGAAGCCGAT AA

```

This corresponds to the amino acid sequence <SEQ ID 636; ORF 154.a>:

```

a154 . pep
1  MTDNSPPPNG HAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP
51  VVTLLMDSAE GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD
101 VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTLL SGSYIAFTPG KSDEAKDVFQ

```



151 VQDIPPVTAI GQSGRLRLNLI GKNDRLNVN SPVLYENFMV GQVESAHFDP  
201 SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI KLNSAPLPAL  
251 LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ  
301 SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS  
351 RLEINADEQS KEHWKQFOT ALNKGLTATI SSNNLLTGSK MIELNDQPSA  
401 SPKLRPHTVY AGDTVIATQG GGLDDLQVKL ADLLDKFDKL PLDKTVAELN  
451 GSLAEKSTL KSANAALSSI DKLVGKPTQ NIPNELNQT KEKPNALIF  
501 SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NSSSKDPIPK  
551 GSR\*

## m154/a154 100.0% identity in 553 aa overlap

	10	20	30	40	50	60
m154.pep	MTDNSPPPN	GHAQARVR	KNNFTLSA	VWLVPLIA	LIAGGWLW	VKEIRNRG
a154	MTDNSPPPN	GHAQARVR	KNNFTLSA	VWLVPLIA	LIAGGWLW	VKEIRNRG
	70	80	90	100	110	120
m154.pep	GIEVNNTVI	KVLSIDVG	RVTRIKLR	DDQKGVET	AQLNADV	SGLIRSDT
a154	GIEVNNTVI	KVLSIDVG	RVTRIKLR	DDQKGVET	AQLNADV	SGLIRSDT
	130	140	150	160	170	180
m154.pep	SGVTGLGT	LLSGSYIA	FTPGKSDE	AKDVQVQD	IPPVTAIG	QSGRLRLN
a154	SGVTGLGT	LLSGSYIA	FTPGKSDE	AKDVQVQD	IPPVTAIG	QSGRLRLN
	190	200	210	220	230	240
m154.pep	SPVLYENFM	VGQVESAH	FDPSDQSV	HYTIFIQSP	NDKLIHSA	SASFWSLE
a154	SPVLYENFM	VGQVESAH	FDPSDQSV	HYTIFIQSP	NDKLIHSA	SASFWSLE
	250	260	270	280	290	300
m154.pep	KLNSAPLP	ALLSGAIS	FDSPKTKN	SKNVKSE	DSFTLYDS	RSEVANLP
a154	KLNSAPLP	ALLSGAIS	FDSPKTKN	SKNVKSE	DSFTLYDS	RSEVANLP
	310	320	330	340	350	360
m154.pep	SVRGLTVG	SPVEYKGL	NVGVVSD	VPYFDRND	SLHLFENG	WIPVRIRI
a154	SVRGLTVG	SPVEYKGL	NVGVVSD	VPYFDRND	SLHLFENG	WIPVRIRI
	370	380	390	400	410	420
m154.pep	KEHWKQFQ	TALNKGLT	ATISSNNL	LTGSKMIE	LNDQPSAS	PKLRPHTV
a154	KEHWKQFQ	TALNKGLT	ATISSNNL	LTGSKMIE	LNDQPSAS	PKLRPHTV
	430	440	450	460	470	480
m154.pep	GGLDDLQV	KLADLLDK	FDKLPLDK	TVAELN	SGSLAEK	STLKSANA
a154	GGLDDLQV	KLADLLDK	FDKLPLDK	TVAELN	SGSLAEK	STLKSANA
	490	500	510	520	530	540
m154.pep	NIPNELNQ	TLKELRTT	LQGVSPQS	PIYGDVQ	NTLQSLDK	TLKDVQPV
a154	NIPNELNQ	TLKELRTT	LQGVSPQS	PIYGDVQ	NTLQSLDK	TLKDVQPV

```

                    550
m154.pep      NSSSKDPIPKGSRX
              |||||
a154          NSSSKDPIPKGSRX
                    550

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 637>:

```

g155.seq
1   atGAAaatcg  GtateCCACG  CGAGTCAtta  tcCGGCGAAA  cccgcgtagc
51  ctgcAcgccc  gCCACCGTTG  CCctgctggg  caAactAGGC  TTTGAAACCG
101 TTGtcgaAAG  CGGTGCAggt  TTGGCGGCAA  GTTTggaCGA  TGCCGCTTAC
151 CAAACAGCAG  GCGCAACCGT  TGCCGACAAA  GCGGCGGTTT  GGGCCTGCCC
201 TTTAATTTAT  AAGGTCAACG  CGCCGTCCGA  AGGCGAGCTG  CCGCTGCTCA
251 AAGAAGGTCA  AACCATCGTC  AGCTTCCTGT  GGCCGCGCCA  AAACGAGGCT
301 TTGGTCGAGG  CCTTGC GCGC  CAAGAAAGTC  AACGCGCTGG  CGATGGACAT
351 GGTTCCTCCG  ATTTCCCGCG  CTCAGGCCTT  GGACGCTTTG  TCTTCAATGG
401 CAAACATCAG  CGGCTACCGC  GCCGTGATTG  AAGCCGCCAA  CGCCTTCGGC
451 CGTTTCTTCA  CCGGTCAAAT  CACTGCCGCC  GGCAAAGTGC  CGCCTGCGCA
501 GGTTTTGGTG  ATTGGCGCCG  GTGTGGCGGG  TTTGGCGGCA  ATCGGTACGG
551 CAAATTCGCT  CGGCGCAGTG  GTGCGCGCGT  TCGATACCCG  CTTGGAAGTG
601 GCGGAACAAA  TCGAATCGAT  GGGCGGTAAG  TTcctGAAAC  TCGACTTCCT
651 GCAAGAATCG  GGCGGCAGCG  GAGACGgcta  CGCCAAAGTG  ATGAGCGACG
701 AATTTATCGC  CGCCGAAATG  AAGCTCTTTG  CCGAACAGGC  GAAAGAAGTG
751 GACATCATCA  TCACCACCGC  CGCCATTCCG  GGCAAACCCG  CTCCCAAGCT
801 GATTACCAAA  GAAATGGTGG  AAAGCATGAA  ATCCGGATCC  GTCATCGTCG
851 ATTTGGCGGC  GACGGGCGGC  AACTGCGAAC  TCACCCGACC  GGGCGAATTG
901 TCCGTAACCG  GCAACGCGCT  GAAAATCATC  GGCTACACCG  ACATGGCAAA
951 CCGCCTTGCC  GGACAGTCTT  CCCAGCTTTA  CGCCACCAAC  TTGGTGAACC
1001 TGACCAAGCT  GTTAAGCCCG  AACAAAGAcg  gcgaAATCAC  GCTGGACTTC
1051 GAAGacgtGA  TTATCCGCAA  TATGACCGTT  ACCCGcgacg  gcgaaATCAC
1101 CTTCCCGCCT  CCGccgaTTc  aggtTTCcgc  ccggccgCAG  CAAACgccgt
1151 ctgaAAAagc  cgcGCCGTGC  GCCAagcccg  AgccGaaacc  tgttCCcctg
1201 tggAAAAaac  tcgCGCCCGC  CGCcatcgCC  GCCGTATTGG  tgctgtgGgt
1251 cggCgcggtc  gcaccgcag  CATCTTGAA  CCACTTTATC  GTCTTCGTCC
1301 TCGCCTGCGT  CATCGGCTAC  CATGTCGTTT  GgaacgTCAG  CCACTCGCTG
1351 CACACACCGC  TGatgtcggt  aaccaaCgcc  atctccGGCA  tcatgggtcgt
1401 cggCGCGCTG  CTGCAAATCG  GTCAGGGcaa  cggcttcgtT  TCgctGCTGT
1451 CGTTTGTTCG  CATCCTGATT  GCCGGCATCA  ATATCTTCGG  CGGCTTTGCG
1501 GTTACACGGC  GTATGCTGAA  TATGTTTAAG  AAAGGGTAA

```

This corresponds to the amino acid sequence <SEQ ID 638; ORF 155.ng>:

```

g155.pep
1   MKIGIPRESL  SGETRVACTP  ATVALLGKLG  FETVVESGAG  LAASLDAAAY
51  QTAGATVADK  AAVWACPLIY  KVNAPSEGEL  PLLKEGQTIY  SFLWPRQNEA
101 LVEALRAKKV  NALAMDMVPR  ISRAQALDAL  SSMANISGYR  AVIEAANAFG
151 RFFTQGITAA  GKVPPAQVLV  IGAGVAGLAA  IGTANSLGAV  VRAFDTRLEV
201 AEQIESMGGK  FLKLDFLQES  GSGDGYAKV  MSDEFIAAEM  KLFAEQAKEV
251 DIIITTAaip  GKPAKPLITK  EMVESMKSGS  VIVDLAATGG  NCELTRPGEL
301 SVTGNVKIi  GYTDMANRLA  GQSSQLYATN  LVNLTKLLSP  NKDGEITLDF
351 EDViiRNMTV  TRDGEITFPF  PPIQVSARPO  QTPSEKAAPA  AKPEPKPVPL
401 WKKLAPAAIA  AVLVLVGAV  APAAFLNHFI  VFVLACVIGY  HVVNVVSHSL
451 HTPLMSVTNA  ISGIMVVGAL  LQIGQGNGFV  SLLSFVAILI  AGINIFGGFA
501 VTRRMLNMFK  KG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 639>:

```

m155.seq
1   ATGAAAATCG  GTATCCACG  CGAGTCATTA  TCCGCGGAAA  CCCGCGTCGC
51  CTGTACGCCC  GCCACCGTCG  CCCTGCTGGG  CAAACTGGGC  TTTGAAACCG
101 TTGTCGAAAG  CGGTGCAAGT  TTGGCGGCAA  GTTTGGACGA  TGCCGCTTAC
151 CAAACAGCAG  GCGCAACCGT  TGCCGACAAA  GCGGCGGTTT  GGGTCTGCCC
201 TTTGATTTAT  AAGGTCAACG  CGCCGTCCGA  ACAGGAAGTG  CCGCTTTTGA
251 ACGAAGGTCA  AACCATCGTC  AGCTTCCTGT  GGCCGCGCCA  AAACGAGGCT

```

```

301 TTGGTCGAAG CCTTGC GCGC CAAGAAAGTG AACGCGCTGG CGATGGATAT
351 GGTGCCCCGC ATTTGCGCGC CGCAGGCTTT GGACGCTTTG TCTTCGATGG
401 CAAACATCAG CGGCTACCGC GCCGTAATTG AAGCCGCCAA CGCCTTCGGC
451 CGTTTCTTCA CCGGTCAAAT TACCGCCGCC GGCAAAGTGC CGCCCGCGCA
501 GGTTTTGGTG ATTGGTGCAG GTGTGGCAGG TTTGGCGGCG ATCGGTACGG
551 CAAACTCGCT CGGCGCAGTG GTACGCGCGT TCGATACCCG CTTGGAAGTG
601 GCGGAACAAA TCGAATCGAT GGGCGGCAAG TTCCTGAAAC TCGACTTCCC
651 ACAAGAATCG GGCGGCAGCG GAGACGGCTA CGCCAAAGTG ATGAGCGACG
701 AATTTATCGC AGCCGAGATG AAGCTCTTTG CCGAGCAGGC GAAAGAAGTG
751 GACATCATCA TCACCACCGC CGCCATTCCG GGCAAACCCG CGCCCAAGCT
801 GATTACCAAA GAAATGGTGG AAAGCATGAA ATCCGGCTCC GTCATCGTCG
851 ATTTGGCGGC GGCGACGGGC GGCAACTGCG AACTCACCCG CCCGGGCGAA
901 TTGTCCGTAA CCGGCAACGG CGTGAAAATC ATCGGCTACA CCGACATGGC
951 AAACCGCCTT GCCGGACAGT CTTCCCAGCT TTACGCCACC AACTTGGTCA
1001 ACCTGACCAA GCTGTTAAGC CCGAACAAAG ACGGCGAAAT CACGTTGGAC
1051 TTCGAAGACG TGATTATCCG CAACATGACC GTTACCCACG ACGGCGAAAT
1101 CACCTTCCCG CCTCCGCCGA TTCAAGTTTC CGCCGAGCCG CAGCAAACGC
1151 CGTCTGAAAA AGCCGTGCCT GCCGCCAAGC CCGAGCCAAA ACCCGTTCCC
1201 CTGTGGAAAA AACTCGCGCC CGCCGTCATC GCCGCCGTCT TGGTACTGTG
1251 GGTGCGCGCG GTCGCACCCG CAGCATTCTT GAACCACTTT ATCGTGTTCG
1301 TTCTCGCCTG CGTCATCGGC TACTACGTCG TCTGGAACGT CAGCCACTCG
1351 CTGCACACAC CGCTGATGTC GGTAACCAAC GCCATCTCCG GCATCATCGT
1401 CGTCGGCGCG CTGCTGCAA TCGGTCAGGG CAACGGCTTC GTTTCGCTGC
1451 TGTCGTTTGT TGCCATCCTG ATTGCCGGA TCAACATCTT CGGCGGCTTT
1501 GCGGTAACAC GCGGTATGCT GAATATGTTT AAGAAAGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 640; ORF 155>:

```

m155.pep
1 MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY
51 QTAGATVADK AAVWVCPLIY KVNAPSEQEL PLLNEGQTI V SFLWPRQNEA
101 LVEALRAKKV NALAMDMPV ISRAQALDAL SSANISGYR AVIEAANAFA
151 RFFTGQITAA GKVPQAQVLV IGAGVAGLAA IGTANSLGAV VRAFDTRLEV
201 AEQIESMGGK FLKLDLPQES GSGDGYAKV MSDEFIAAEM KLF AEQAEV
251 DIIITTAIIP GKPAKLIK EMVESMKSGS VIVDLAAATG GNCELTRPGE
301 LSVTGNVSKI IGYTDMANRL AGQSSQLYAT NLVNLTKLLS PNKDGEITLD
351 FEDVIIRNMT VTHDGEITFP PPPIQVSAQP QQTPEKAVP AAKPEPKVP
401 LWKKLAPAVI AAVLVLVWGA VAPAAFLNHF IVFVLACVIG YYVWVNVSHS
451 LHTPLMSVTN AISGIIVVGA LLQIGQGNF VSLLSFVAIL IAGINIFGGF
501 AVTRMLNMF KKG*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 155 shows 97.9% identity over a 513 aa overlap with a predicted ORF (ORF 155.ng) from *N. gonorrhoeae*:

```

m155 / g155 97.9% identity in 513 aa overlap
              10      20      30      40      50      60
m155.pep     MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAY QTAGATVADK
              |||||
g155          MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAY QTAGATVADK
              10      20      30      40      50      60

              70      80      90      100     110     120
m155.pep     AAVWVCPLIYKVNAPSEQEL PLLNEGQTI V SFLWPRQNEALVEALRAKKV NALAMDMPV
              ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g155          AAVWACPLIYKVNAPSEGELE PLLKEGQTI V SFLWPRQNEALVEALRAKKV NALAMDMPV
              70      80      90      100     110     120

              130     140     150     160     170     180
m155.pep     ISRAQALDALSSMANISGYRAVIEAANAFA GRFFTGQITAAGKVPPAQVLVIGAGVAGLAA
              |||||
g155          ISRAQALDALSSMANISGYRAVIEAANAFA GRFFTGQITAAGKVPPAQVLVIGAGVAGLAA

```

	130	140	150	160	170	180
m155.pep	190	200	210	220	230	240
	IGTANSLGAVVRAFDTRLEVAEQIESMGGKFLKLDLPQESGGSGDGYAKVMSDEFIAAEM					
g155	IGTANSLGAVVRAFDTRLEVAEQIESMGGKFLKLDLPQESGGSGDGYAKVMSDEFIAAEM					
	190	200	210	220	230	240
m155.pep	250	260	270	280	290	300
	KLFAEQAKEVDIIITTAIPGKPAPKLITKEMVESMKSGSVIVDLAAATGGNCELTRPGE					
g155	KLFAEQAKEVDIIITTAIPGKPAPKLITKEMVESMKSGSVIVDLAA-TGGNCELTRPGE					
	250	260	270	280	290	
m155.pep	310	320	330	340	350	360
	LSVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT					
g155	LSVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT					
	300	310	320	330	340	350
m155.pep	370	380	390	400	410	420
	VTHDGEITFPPPIQVSAQPQOTPSEKAVPAKPEPKPVPLWKKLAPAVIAAVLVLWVGA					
g155	VTRDGEITFPPPIQVSARPQOTPSEKAAPAKPEPKPVPLWKKLAPAAIAAVLVLWVGA					
	360	370	380	390	400	410
m155.pep	430	440	450	460	470	480
	VAPAAFLNHFIVFVLACVIGYVWVNVSHSLHTPLMSVTNAISGIIIVVGALLQIGQGNGF					
g155	VAPAAFLNHFIVFVLACVIGYHVWVNVSHSLHTPLMSVTNAISGIMVVGALLQIGQGNGF					
	420	430	440	450	460	470
m155.pep	490	500	510			
	VSLLSFVAILIAGINIFGGFAVTRMLNMFKKGX					
g155	VSLLSFVAILIAGINIFGGFAVTRMLNMFKKGX					
	480	490	500	510		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 641>:

```

a155.seq
1   ATGAAAATCG GTATCCCACG TGAGTCATTA TCCGGCGAAA CCCGCGTCGC
51  CTGTACGCCC GCCACCGTCG CCCTGCTGGG CAAACTGGGC TTTGAAACCG
101 TTGTCGAAAG CGGCGCAGGT TTGGCGGCAA GTTTGGACGA TGCCGCTTAC
151 CAAGCAGCAG GCGCAACCGT TGCCGACAAA GCAGCGGTTT GGGCATACCC
201 TTTAATTAT AAGGTTAACG CGCCGTCCGA AGACGAGCTG CCGTGTCTCA
251 AAGAAGGACA GACCATCGTC AGCTTCCTGT GGCCGCGCCA AAACGAGGCT
301 TTGGTCGAAG CCTTGCGCGC CAAGAAAGTG AACGCGCTGG CAATGGACAT
351 GGTGCCCCGC ATTTGCGCGC CGCAGGCTTT GGACGNTTGT TCTTNGATGG
401 CAAACATCAG CGGCTACCGC GCCGTGATTG AAGCCGCCAA CGCCTTCGGC
451 CGTTTNTTCA CCGGCCAAAT TACTGCCGCA GGCAAAGTGC CGCCGCGCA
501 GGTTTTGGTG ATTGGTGCAG GTGTGGCAGG TTTGGCGGCG ATCGGTACGG
551 CAAACTCGCT CGGCGCAGTG GTACGCGTGT TCGATACCCG CCTG.AAGTG
601 GCGGAACAAT TAGAATCGAT GGGCGGCAAG TTCTTGAAAC TCGACTTCCC
651 GCAAGAATCG GGCGGCAGCG GCGACGGCTA CGCAAAGTG ATGAGCGACG
701 AATTTATCGC CGCCGAGATG AAGCTTTTGT CCGAGCAGGC GAAAGAAGTG
751 GACATCATCA TCACCACCGC CGCATTCCG GGCAAACCCG CGCCCAAGCN
801 NNTNANCAA GAAATGGTCG AAAGCATGAA ACCCGGCTCC GTCATCGTCG
851 ATTTGGCGGC GCGACGGGC GGCAACTGCG AACTCACCAA ACAGGGCGAA
901 TTGTTCTGTA CCGGCAACGG CGTGAAATC ATCGGCTACA CCGACATGGC
951 AAACCGCCTT GCCGGACAGT CTTCGAGCT TTACGCCACC AACTTGGTCA
1001 ACCTGACCAA GCTGTTAAGC CCGAACAAAG ACGGCGAAAT CACGCTGGAC
1051 TTCGAAGACG TGATTATCCG CAACATGACC GTTACCCGCG ACGGCGAAAT
1101 CACCTTCCCG CCTCCGCCGA TTCAAGTTT CGCCCAACCG CAGCAAACCG
1151 CGTCTGAAAA AGCCGCGCCT GCCGCAAGC CCGAACCGAA ACCCGTTCCC

```

```
1201 CTGTGGAAAA AACTCGCGCC CGCCNTNATC GCCGCCGTGT TGGTACTGTG
1251 GGTCGGCGCG GTCGCACCCG CAGCATTCTT GAACCACTTT ATCGTCTTCG
1301 TCCTCGCCTG CGTCATCGGC TACTATGTCG TTTGGAACGT CAGCCACTCG
1351 CTGCACACAC CGCTGATGTC GGTGACCAAC GCCATTTCCG GCATCATCGT
1401 CGTCGGCGCG CTGCTGCAAA TCGGTCAGGG CAACGGCTTC GTTTCGCTGC
1451 TGTCGTTTGT TGCCATCCTG ATTGCCAGCA TCAACATCTT CGCGCGCTTC
1501 TTTGTAACGC GCGGATGCT GAATATGTTT AGGAAAGGGT AA
```

This corresponds to the amino acid sequence <SEQ ID 642; ORF 155.a>:

```
a155.pep
1  MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY
51  QAAGATVADK AAVWAYPLIY KVNAPSEDEL PLLKEGQTIV SFLWPRQNEA
101 LVEALRAKKV NALAMDMPVR ISRAQALDXL SXMANISGYR AVIEAANAFG
151 RXFTGQITAA GKVPQAQVLV IGAGVAGLAA IGTANSLGAV VRVFDTRLXV
201 AEQLESMGGK FLKLDFFQES GSGDGYAKV MSDEFIAAEM KLFAEQAKEV
251 DIIITTAaip GKPAKXXXK EMVESMKPGS VIVDLAAATG GNCELTKQGE
301 LFVTGNGVKI IGYTDMANRL AGQSSQLYAT NLVNLTKLLS PNKDGEITLD
351 FEDVIIRNMT VTRDGEITFP PPPIQVSAQP QOTPSEKAAP AAKPEPKPVP
401 LWKKLAPAXI AAVLVLVWGA VAPAAFLNHF IVFVLACVIG YYVWVNVSHS
451 LHPTPLMSVTN AISGIIIVGA LLQIGQGNF VSLLSFVAIL IASINIFGGF
501 FVTRRMLNMF RKG*

m155/a155 95.3% identity in 513 aa overlap

      10      20      30      40      50      60
m155.pep  MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAYQTAGATVADK
a155      MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAYQAAGATVADK
      10      20      30      40      50      60

      70      80      90     100     110     120
m155.pep  AAVWVCPLIYKVNAPSEQELPLLNEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMPVR
a155      AAVWAYPLIYKVNAPSEDELPLLKEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMPVR
      70      80      90     100     110     120

      130     140     150     160     170     180
m155.pep  ISRAQALDALSSMANISGYRAVIEAANAFGRFFTQGITAAAGKVPQAQVLVIGAGVAGLAA
a155      ISRAQALDXLSXMANISGYRAVIEAANAFGRXFTQGITAAAGKVPQAQVLVIGAGVAGLAA
      130     140     150     160     170     180

      190     200     210     220     230     240
m155.pep  IGTANSLGAVVRAFDTRLEVAEQIESMGKFLKLDFFQESGSGDGYAKVMSDEFIAAEM
a155      IGTANSLGAVRVFDTRLXVAEQLESMGGKFLKLDFFQESGSGDGYAKVMSDEFIAAEM
      190     200     210     220     230     240

      250     260     270     280     290     300
m155.pep  KLFAEQAKEVDIIITTAaipGKPAKKLITKEMVESMKSGSVIVDLAAATGGNCELTRPGE
a155      KLFAEQAKEVDIIITTAaipGKPAKXXXKEMVESMKPGSVIVDLAAATGGNCELTKQGE
      250     260     270     280     290     300

      310     320     330     340     350     360
m155.pep  LSVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT
a155      LFVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT
      310     320     330     340     350     360

      370     380     390     400     410     420
m155.pep  VTHDGEITFPPPIQVSAQPQOTPSEKAVPAKPEPKPVPLWKKLAPAVIAAVLVLVWGA
a155      VTRDGEITFPPPIQVSAQPQOTPSEKAAPAAKPEPKPVPLWKKLAPAXIAAVLVLVWGA
      370     380     390     400     410     420
```

```

      430      440      450      460      470      480
m155.pep  VAPAAFLNHFIVFVLACVIGYYVWVNVSHSLHTPLMSVTNAISGIIIVVGALLQIGQGNGF
          |||||
a155      VAPAAFLNHFIVFVLACVIGYYVWVNVSHSLHTPLMSVTNAISGIIIVVGALLQIGQGNGF
          |||||
      430      440      450      460      470      480

      490      500      510
m155.pep  VSLLSFVAILIAGINIFGGFAVTRRMLNMFKKGX
          |||||:|||||
a155      VSLLSFVAILIASINIFGGFFVTRRMLNMFKKGX
          |||||
      490      500      510

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 643>:

```

g156.seq
1  ATGACTTTCG CCTATTGGTG CATTCTGATT GCCTGCCTAT TGCCGCTTTT
51  TTGTGCGGCG TATGCCAAAA AAGCGGGCGG ATTCCGGTTT AAAGACAACC
101 ACAATCCTCG CGGTTTTCTG GCACATACGC AAGGCGCAGC CGCCCGTGCC
151 CACGCCGCGC AGCAAAACGG TTTTGAAGCC TTGCACCGT TTGCCGCCGC
201 CGTTTTGACG GCACACGCAA CCGCAATGC CGGACAAGCA ACCGTCAACA
251 CGCTTGCCGG ATTGTTTCATC CTGTTCCGCC TCGCCTTTAT CTGGTGCTAC
301 ATCGCAGACA AAGCAGCATT GCGCTCGCTG ATGTGGGCGG GCGGATTTGC
351 CTGCACCGTC GGACTGTTTG TCGCGGCTGC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 644; ORF 156.ng>:

```

g156.pep
1  MTFAYWCILI ACLLPFCAA YAKKAGGERF KDNHNPRGFL AHTQGAAARA
51  HAAQONGFEA FAPFAAAVLT AHATGNAGQA TVNTLAGLFI LFRLAFIWCY
101 IADKAALRSL MWAGGFACTV GLFVAAA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 645>:

```

m156.seq
1  ATGACTTTCG CCTATTGGTG TATTCTGATT GCCTGCCTAT TGCCGCTTTT
51  TTGTGCGGCG TATGCCAAAA AAGCGGGCGG ATTCCGGTTT AAAGACAACC
101 ACAATCCTCG CGGTTTTCTA GCGCACACGC AAGGCGCAGC CGCCCGTGCC
151 CACGCCGCGC AGCAAAACGG TTTTGAAGCC TTGCACCGT TTGCCGCCGC
201 CGTTTTGACG GCACACGCAA CCGCAATGC GCGCAATCG ACCATCAACA
251 CGCTTGCTG CCTGTTTCATC CTGTTCCGCC TCGCCTTTAT CTGGTGCTAT
301 ATCGCCGACA AAGCCGCTAT GCGCTCACTG ATGTGGGCAG GCGGATTTGC
351 CTGCACCGTC GGGCTGTTTG TCGCGGCTGC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 646; ORF 156>:

```

m156.pep
1  MTFAYWCILI ACLLPFCAA YAKKAGGERF KDNHNPRGFL AHTQGAAARA
51  HAAQONGFEA FAPFAAAVLT AHATGNAAQS TINTLACLFI LFRLAFIWCY
101 IADKAAMRSL MWAGGFACTV GLFVAAA*

```

Computer analysis of this amino acid sequence gave the following results:  
Homology with a predicted ORF from *N. gonorrhoeae*

m156 / g156 96.1% identity in 127 aa overlap

```

      10      20      30      40      50      60
m156.pep  MTFAYWCILIACLLPFCAAYAKKAGGERFKDNHNPRGFLAHTQGAAARAHAAQONGFEA
          |||||
g156      MTFAYWCILIACLLPFCAAYAKKAGGERFKDNHNPRGFLAHTQGAAARAHAAQONGFEA
          |||||
      10      20      30      40      50      60

      70      80      90     100     110     120
m156.pep  FAPFAAAVLTAAHATGNAAQSTINTLACLFILFRLAFIWCYIADKAAMRSLMWAGGFACTV
          |||||:|||||

```